STUDIES OF CUTANEOUS PAIN: A SURVEY OF RESEARCH SINCE 1940

BY K. R. L. HALL

Department of Psychology, Barrow Hospital, Bristol

Introduction (pp. 279–280). II. *Normal pain threshold* (pp. 280–286): (i) *Uniformity of threshold* p. 280–282); (ii) *Variability: attitude factors* (pp. 282–284); (iii) *Reactivity to pain* (pp. 284–286). I. *Pain response in psychiatric conditions* (pp. 287–291): (i) *Pain thresholds in psychoneurotics* (pp. 287–8); (ii) *Response to pain in schizophrenia and depression* (p. 289); (iii) *Effects of lobotomy* (pp. 289–291). . *Conclusions* (pp. 291–292). *References* (pp. 292–294).

I. INTRODUCTION

esearch on cutaneous pain was, until 1940, concerned mainly in establishing the ana- mical and physiological basis for the peripheral reception and transmission of stimuli the central nervous system. A very thorough review of this research between 1930 and 40 has been made by Stone & Jenkins (1940), and brief accounts of the present state knowledge of the 'pain sense' are contained in Morgan & Stellar (1950) and in Jenkins)51).

Beyond this sensory research on cutaneous pain, very little seems to have been done, ior to 1940, to establish experimentally the more complex conditions involved in pain a discriminatory threshold phenomenon, as a perceptual process, or in its emotional ntext in human subjects. It is primarily these more central aspects of pain which have en investigated in some detail, both in normal and abnormal conditions, since the scription by Hardy, Wolff & Goodell (1940) of apparatus and techniques well adapted the problem. The purpose of the present survey, therefore, is to give an account and preciation of the considerable amount of work of psychological interest which has been ried out since that date.

A few fairly comprehensive accounts of pain research, mainly on the anatomical, ysiological and medical aspects of the subject, have appeared in the last 12 years, tably those of Lewis (1942), Livingston (1943), Wolff & Hardy (1947), and Wolff & olf (1948). The review by Wolff & Hardy gives a valuable summary of physiological earch, much of which is also discussed by Wolff & Wolf. A survey by Edwards (1950) 'Recent research on pain perception' seems, however, to be the only detailed contem- rary account written from the perspective of the psychologist. Edwards's outline deals efly with four related problems:

1) A definition of pain.
2) Psychophysics of pain.
3) Peripheral-nervous mechanisms involved in pain perception.
4) Role of the central nervous system in certain abnormalities of pain perception.

n our present survey, we shall overlap very little with his account. With regard to the estion of a definition for pain, there is, it seems, no advantage in listing the many sibilities. It has been suggested by Dunbar (1947) that one reason for the confusion in ard to the subject of pain has been the failure of investigators to define what they mean pain. To the present writer, however, it seems that what is meant by pain should be

apparent, in each investigation, from the description of the experimental conditions and controls, the instructions, the results, and the conclusions. The reason, therefore, for the occasional 'confusion', by which she may perhaps mean 'discrepancy' between results we shall see to be mainly due to some inadequacy in description and in control of conditions. 'Pain' may be studied as a 'sensation' in one experiment, as a 'perception' and as involving 'attitudes' in another, and as related to emotional behaviour in a third.

II. Normal pain threshold

The problem of measuring the threshold for pain in normal subjects has been very thoroughly investigated since the first paper on the subject by Hardy *et al.* (1940) in which they reintroduced thermal methods for producing cutaneous pain. The work can broadly be classified under three main headings:

(i) Investigations showing conditions for the uniformity of the threshold in man.

(ii) Investigations which emphasize the variability of threshold, particularly where psychological factors of attitude are influential.

(iii) Investigations of reactivity to pain.

(i) *Uniformity of threshold*

Dealing first with the investigations that come under heading (i), Hardy *et al.* began by experimenting on themselves, using a radiant heat apparatus, the light from which was focused on to the forehead, and exposed for 3 sec./min. The heat intensity required to produce pain was measured with a radiometer placed at the aperture, the threshold being expressed in millicalories per second per square centimetre (m.cal./sec./cm.^{2}). All threshold values quoted in the present paper will be given in this unit of measurement except where otherwise stated.

The authors had decided to use a thermal method for producing pain, because they considered it to have several advantages over other methods, such as prick-pressure amongst which they mention the following:

(1) The intensity of the stimulating agent can be precisely measured.

(2) The sensory threshold for pain obtained from a thermal stimulus is said to be a sharply defined experience thus allowing for a higher degree of accuracy in threshold determination than is available with other methods.

(3) The method is flexible in that time of exposure to the stimulus, area and state of the skin, etc., can be varied at will.

In experimenting on themselves over a period of a year, the authors reported a high degree of uniformity in the pain threshold, and also a very close agreement amongst themselves, all three having almost exactly the same threshold. In understanding this uniformity, it is important to take into account very carefully the exact experimental procedure and instructions used. The threshold was obtained by increasing the stimulus intensity step by step until the pain perception point was reported. This threshold (P.P.T.) is said by the same authors (1943) to be easily recognizable by untrained but instructed subjects. The sensation was that of heat finally 'swelling' to a distinct, sharp stab of pain at the end of the 3 sec. exposure. Provided, therefore, that a subject is capable of maintaining an 'objective' attitude of adequate concentration on the stimulus, the threshold is said to vary extremely little.

The same conclusion was reached by Schumacher, Goodell, Hardy & Wolff (1940) and

Hardy *et al.* (1943) after applying their technique and method of instruction to 200 normal subjects. Furthermore, these writers conclude that the threshold values so obtained are '. . . independent of age, sex, various emotional states and fatigue within the limits of the subject to maintain an unprejudiced detached attitude' (p. 5). Their subjects were repeatedly examined at intervals of several days and weeks, the distribution of P.P.T.'s being described as bell-shaped with a limit of deviation from the mean value 226 of -16 to $+13\%$. The intra-group variation was said to be no greater than the intra-individual variation for the four authors over a period of 24 hr. without sleep.

One further point that was studied as incidental to the main investigation was the subject's personal estimation of his own pain-sensitivity. With few exceptions, it is said that no relation was found between such opinions and the pain threshold as measured here.

From these studies, the evidence for the uniformity of threshold is convincing, but only in so far as it relates to definitely instructed subjects who are able to maintain a detached, objective attitude. The 'stimulus sensation' that is to equal 'pain' is carefully defined, thus all the usual perceptual context of 'pain' is, so far as possible, eliminated.

Physiological factors

The possible effect on the heat-pain threshold of certain physiological factors has been carefully studied and is summarized below.

(1) Effect of skin temperature: Hardy, Goodell & Wolff (1951) found that, where the forehead was used as stimulus area, there is no likelihood of error in the threshold measurements provided that the laboratory temperature is maintained between 20 and 30° C., because the skin temperature of this area remains constant. For thresholds obtained on other skin areas, such as the forearm, they suggest a correction formula, the validity of which is confirmed by Wertheimer & Ward (1952).

(2) Size of area of skin stimulated (Hardy *et al.*, 1943). This does not affect threshold, thus indicating that there is no spatial summation for pain.

(3) Profuse sweating on the stimulus area (Chapman & Jones, 1944) does not affect threshold.

(4) Moderate hypercapnia produced by breathing mixtures of 5 and 7·5% CO_2 elevated threshold significantly in all subjects (Chapman & Smith, 1948), whereas hypoxia was found to affect pain sensation very much less than other sensations, such as visual. This analgesic effect of inhaled CO_2 was found to be chiefly of central origin and not peripheral.

(5) The effect of cutaneous hyperalgesia (area of burn on the forearm) was to cause a more intense pain when the stimulus was applied to this area than when applied to control areas of skin (Hardy, Wolff & Goodell, 1950).

(6) Administration of opiates, such as morphine, considerably raised the P.P.T. (Wolff, Hardy & Goodell, 1940), the highest threshold-raising effect of which the drug was capable being 100% above the control threshold about 90 min. after administration.

The dol scale

Thus, certain physiological factors are able, as was to be expected, to affect the threshold uniformity in normals, but, in other respects, this uniformity was sufficiently convincing to lead Hardy, Wolff & Goodell (1947) to try to devise a scale of pain intensity based on

discrimination of differences in pain sensation. Previously, Herget, Granath & Hardy (1941) had obtained j.n.d.'s[1] for the whole range of thermal stimulation from just perceptibly warm, up to the onset of pain, giving altogether about thirty steps of distinguishable thermal sensation. Hardy *et al.*, using an improved radiant-heat apparatus with a 500 W. bulb as heat-source, tried to establish the j.n.d.'s for the whole range of painful stimulation, with themselves as subjects, beginning from the lowest point perceptible as pricking pain (at 220 m.cal./sec./cm.^{2}) up to 680—a point at which definite tissue damage was caused. There was no apparent increase in the value of ΔI between the threshold stimulus and 290, but between 290 and 340, ΔI more than doubled. Between 340 and 400, there was little change but, about 400, there was a final sharp rise in the ΔI values. Thus, although 420 and 480 could be distinguished, 480 and 580 could not, and it was not until the stimulus was increased to 680 that a j.n.d. could be observed, while discrimination for stimuli greater than 680 was impossible so that this measure came to be described as a 'ceiling' pain.

The writers found, however, that, except under the best possible experimental circumstances, it was impossible to distinguish stimuli closer together than two j.n.d.'s; hence the arbitrary unit for the pain sensation scale was chosen as equivalent to two such steps = 1 dol (unit of pain), giving a total scale of 10 dols, covering the intensity range 220–680 m.cal./sec./cm.^{2}.

In a follow-up study, the same writers (1948) obtained results by 'fractionation' of a stimulus in terms of an initial standard (for example, reporting it to be one-quarter of the standard) which were said to confirm their establishment of a scale of painfulness derived by integration of j.n.d.'s. The implication is that, for cutaneous pain sense, the sensory magnitude of j.n.d.'s is the same in all parts of the scale, so that subjective estimates by normal adults of degrees of painfulness should be surprisingly accurate, even without prior experience of such estimates. Consequently, they consider that their dol scale should be useful for all kinds of comparative investigations of pain appreciation, as, for example, the study of pain intensity experienced in childbirth (Hardy & Javert, 1949).

In the work so far described, the subjects have been highly sophisticated and instructed in regard to the general experimental procedure and as to the actual nature and quality of the sensation which they were required to describe as painful. Nothing, so far as possible, is left to individual interpretation, to attitude factors which might well be crucial in the psychological experience of pain for naïve, uninstructed subjects. The critical difference emerging from the work now to be described is the much greater range of variability in normal threshold and the clearly demonstrable effect of attitude on pain appreciation.

(ii) *Variability: attitude factors*

This greater variability in normal threshold was particularly emphasized by Lanier (1943) and Chapman & Jones (1944).

Lanier used an electric current stimulator for obtaining pain thresholds on fifteen college women. The experiment was repeated on a second day with the same subjects. The thresholds, obtained by the method of minimal changes, show a far greater relative variability than those of Schumacher *et al.* (1940); being 49% the first day and 45% the second. These indices are almost 50 times greater than the value reported for thermal stimulation. He further found, from calculating reliability coefficients, that the pain

[1] Just noticeable differences.

threshold of an individual may vary considerably from day to day and from one skin area to another, this variability being very marked in some subjects, and much less so in others of his group.

Chapman & Jones found that the P.P.T.'s for a group of 200 normal subjects of both sexes, age range 10–85, and including both Europeans and Negroes, showed a far greater intra-group deviation than that found by Schumacher *et al.*, the variation being from −40 to +50% around the mean value. Furthermore, P.P.T. was found to increase with age, while there was a tendency for male subjects to have higher thresholds than the female. The statistical evaluation of these findings is, however, insufficient for any definite conclusion to be drawn from them.

Kennard (1952) (using more adequate statistical checks) also reported a sex difference, age groups being equivalent, in both normals and patients, and she was able to eliminate one factor—skin thickness—as a possible explanation for the difference. She is inclined to think that it is due to the difference in the culturally expected response to pain of the two sexes, Schumacher *et al.* having failed to observe it because of the degree of training and sophistication in their subjects. Another factor, said by Chapman & Jones to contribute to the intra-group variability, is a racial difference as between Negro and European subjects. For eighteen Negroes, P.P.T. was found to be lower than the values for eighteen Europeans of the same age and sex. They also report that three other factors—mental fatigue after 8 hr. study period, nervous tension and diurnal variation—seemed to cause some variation in threshold in some of their subjects.

The main finding as to the influence of age seems an important, and probably an expected, one. Dunbar (1947), in discussing this, suggests that we still do not know whether it is due to a difference in physiological threshold or to a difference in the response of the total personality—for example, a greater inhibition of response in the older groups. These two ways of expressing the problem are complementary to one another, and not exclusive, and it seems that detailed experimentation on possible relevant factors, such as length of stimulus exposure time, could contribute much on this question.

Confirmation of these findings of a wide range of normal variation came from Clausen & King (1950) where the standard apparatus and instructions were used on thirty-one untrained subjects (students), age range 19–34. Even within this restricted age group, and in spite of the use of instruction as to the nature of the pain end-point, the threshold variation was ±22·9 S.D. Trying to account for the variation, they suggest that 'educational level' may be an important factor 'because the introspective nature of the task yields a different pattern of reporting in those skilled in the exact use of descriptive words and those whose descriptions are more sparse and dependent upon analogy' (p. 304).

This brings us to a general consideration of those psychological factors, conveniently classed under the term 'attitude', which seem to account both for the uniformity of results in one set of investigations and the relative variability of thresholds in the other.

Chapman & Jones themselves pointed out (1944) that 'the variable that seemed the most important, and the most difficult to standardize, was the subject's description of the pain-perception end-point' (p. 89). Such descriptions were considered of more value when the experimenter made no comments prior to the performance of the test. In other words, their subjects were left relatively free to interpret the situation and the required

response to it in accordance with their own personal attitudes whether specific to the task and the experiment or general towards life experiences of pain. The importance of the instructional 'set', although it is recognized by the writers, is never adequately heeded in the actual experimental procedures. So far as the present writer can discover, there is only one paper, by Chapman, Cohen & Cobb (1946), in which some attempt is made to investigate the effect of varying the form of instruction given in two successive tests, the subjects being groups of neurotic patients. In the first test, the instruction was 'free' and no detailed description as to the nature of the sensation was given. In the second test, they were given the standard instructions to look for a sharp, jabbing, swelling end-point, etc. Results for the two tests were said to be almost identical, but, without a more controlled procedure and adequate statistical treatment of the test-retest data, it seems impossible to reach any conclusion from this finding.

The importance of attitude is clearly indicated in the careful work of Rey (1943) who found a very much wider threshold scatter for his algesimetric results than that given by the same subjects for touch threshold. He attributes this difference to the attitude change that takes place as soon as a subject knows he is going to take part in a pain threshold test.

So far we have discussed attitude as derived from the actual experimental situation and form of instruction given to non-sophisticated subjects. Several attempts have also been made to investigate the effect of experimental variations, such as distraction, suggestion, and hypnosis upon the P.P.T. Wolff & Goodell (1943) reported the following results:

(1) Effect of repeating digits forwards and backwards raised P.P.T. 45%.

(2) Effect of distraction from stories of a highly adventurous nature read aloud raised it 16%.

(3) Effect of autosuggestion 'I won't feel pain' raised it 20%.

(4) Shallow hypnosis raised it 40%.

(5) A clanging noise raised it 38%.

(6) Effect of giving 'dummy' aspirin tablet, after first giving ordinary aspirin, raised it 20% in one subject but had no effect on another.

Ischlondsky (1949), using a prick-pressure apparatus, obtained similar results by imposing various conditions of distraction, visual, acoustic, intellectual or motor, while West, Neill & Hardy (1952) found that hypnotic suggestions of anaesthesia will alter the P.P.T. considerably, there being a general correlation between the depth of trance and the degree to which P.P.T. can be altered.

(iii) *Reactivity to pain*

In addition to the reported perceptual threshold for pain, Chapman & Jones (1944) also obtained a Pain Reaction Threshold (P.R.T.) from their normal subjects, this being defined as the first objective evidence of wincing or withdrawal from the painful stimulus. After the P.P.T. had been reached, the stimulus intensity was further increased until such evidence of reaction was observed.

Although this seems to have been the first occasion on which reaction to, as distinct from perception of, heat pain was specifically studied, physiological reactions to cold pain had been investigated by Hines & Brown (1932)—quoted in Wolf & Hardy (1941)—and by Wolf & Hardy (1941). The pain stimulus was immersion of the hand in very cold water, and the reaction produced by this was a rise in the blood pressure in proportion both to

the intensity of pain experienced and to the degree of cold. When Hines & Brown originally used this 'cold-pressor' test, they divided their normal subjects, on the basis of their results, into three groups according to the amount of rise in blood pressure occurring in the test.

A similar classification was made by Chapman & Jones (1944) of their 200 normal subjects. In group (1) (most reactive), the subject reacted, by wincing or withdrawal, at the same point as he first perceived pain—this was described as a 'fusion' of P.P.T. and P.R.T. In group (2) (average reactive), there was some displacement between P.P.T. and P.R.T., while in group (3) (least reactive), there was a wide displacement (between 50 and 70%) between the two thresholds.

Observations of this kind should help in an understanding of the overall response the individual is making to the pain situation, the degree of discrepancy between the two thresholds having some possible significance in terms of attitudes of inhibition or emotionality, as is indicated by later work on psychiatric patients.

Table 1. *Uniformity—variability of* P.P.T.'s *in normal subjects expressed in terms of the coefficient of variation*

Authors	No. of subjects	Sex	Age range (years)	Coefficient of variation (%)
Schumacher *et al.* (1940)	150	Both	Not stated	± 1
Chapman *et al.* (1947)	56	Both	18–45	± 9·8
Clausen & King (1950)	31	Both	19–34	±12·6
Chapman & Jones (1944)	200	Both	10–85	±14·8
Lanier (1943) (electrical stimulation)	15	F.	Not stated	(1) ±49·0 (2) ±45·0

Other forms of reaction to pain in normal subjects that have been studied are Ischlondsky's (1949) technique for registering the amount of pupillary dilation, and measurement of the G.S.R. accompanying pain (Furer & Hardy, 1950). The latter studied heat-pain response in four normal adults daily for 6 months, and recorded resistance changes during the experimental session. They found that the drop in resistance following a painful stimulus was very rapid, but their main finding was the apparent adaptation to the pain situation that occurred markedly in two subjects after repeated exposure to the painful stimuli.

From the investigations on normal pain threshold reviewed in this section, the chief point of interest has been the evidence for uniformity, on the one hand, and for variability on the other. This evidence is summarized in Table 1, the degree of variation in each set of results being expressed in terms of the coefficient of variation. Where this was not stated in the original papers, it has been calculated by the present writer.

Discussing Lanier's findings in relation to those of Schumacher *et al.*, Jenkins (1951) suggests that the discrepancy might be related in some way to the two sets of cutaneous pain fibres, the one myelinated and moderately rapid in conduction, the other unmyelinated and slow. Bigelow, Harrison, Goodell & Wolff (1945) have shown that invariably, on normally innervated skin, the threshold for 'burning' pain is lower than that for 'pricking' pain, this being apparently due to the difference in conduction rate of the two forms of sensation. If, therefore, the nature of the sensation to be experienced was not accurately defined, subjects would be uncertain as to which end-point to respond.

It is very doubtful whether any peripheral explanation such as that of Jenkins is adequate to account for the extent of the reported discrepancies, nor is it likely, as has been suggested by Edwards (1950), that difference in the method used for obtaining a threshold would account for much of the variation in view of the fact that almost all investigators followed the same general procedure. Most of the possible physiological variables have, as we have seen, been carefully investigated, but, even where these are adequately controlled, there remain considerable differences due to what can best be termed provisionally 'attitude factors'.

Under this heading, it is convenient to include at present all the following:

(1) Effect of ageing.

(2) Sex and racial difference.

(3) Effect of instructional 'set'.

(4) Differences due to verbal intellectual level.

(5) Differences due to emotional stability or instability, and other personality factors.

It is necessary briefly to comment on these items in order to appreciate both the positive achievement so far made and the many possibilities for psychological research on pain which remain for investigation. First, it should be possible to determine experimentally whether the observed differences between age groups, sexes and racial groups are due primarily to differences in central attitude (in terms of learning mechanisms, to differences in inhibition, excitation, and so on) rather than to physiological differences of a peripheral kind. In view of the possibility of achieving uniformity of threshold demonstrated by Schumacher *et al.*, the hypothesis that central factors are crucial is the more likely. Secondly, with regard to instructional and other forms of objectively induced 'set', it should obviously be possible systematically to vary the conditions and context of the 'pain' situation so as to estimate with some accuracy how far variation in pain response may be due to anticipation, inertia, and other individual differences in the subjects. An attempt at experimental control of these variations has recently been completed on psychiatric patients by Hall & Stride (1953), and could usefully be extended to the normal field. 'Objectivity' of reporting the pain sensation has been convincingly demonstrated, but the subjective factors in pain appreciation have scarcely been investigated with any thoroughness. Thirdly, differences due to personality factors, such as introspective intellectual approach, emotionality and so on, possibly having a constitutional basis, can scarcely be disregarded. Sheldon & Stevens (1942), for example, suggest that the person of mesomorphic physique shows typically a Spartan indifference to pain, physiological insensitivity and lack of advance apprehension of painful stimuli, whereas the ectomorph is said to be hypersensitive to pain, with a tendency to overrespond. So far, there does not seem to be any objective confirmation of these observations, two studies of pain resistance and self-ratings for pain tolerance on male university students (Janoff, Beck & Child, 1950; Child, 1950) showing only very low correlations with the Sheldon constitutional components. It is possible, however, that the very homogeneity of the samples used in these investigations has obscured any positive relationships, it being probable that all such 'volunteer' subjects would be highly motivated to 'do well', i.e. to show high resistance in the pain situation.

III. Pain response in psychiatric conditions

The effect of abnormal mental state upon response to pain stimulation has been mainly studied on psychoneurotic patients, although a few investigations have been reported on schizophrenics and depressions and on the effects of lobotomy. The work on pain in psychiatric conditions will, therefore, be reviewed under three headings:

(i) Psychoneurotic response.

(ii) Response in schizophrenic and depressive patients.

(iii) Effects of lobotomy.

(i) *Pain thresholds in psychoneurotics*

The first study on pain in psychoneurotics, in which the radiant heat apparatus and procedures already described were used, was reported by Chapman (1944). A fuller version of his results was made by this author with Finesinger, Jones & Cobb (1947). The most consistent finding emerging from this and the following studies is the much earlier reaction point of the neurotics than that of comparable normals. For example, in the second paper, thresholds for fifty neurotics and for fifty-six normals are given as:

	P.P.T.	P.R.T.
Neurotics	288	314
Normals	283	347

It is stated that the difference between the P.R.T. values is statistically significant, and that, whereas the P.P.T.'s for normals seem to follow a normal distribution curve, those for the patients show a slightly greater range with a bimodal tendency. There seems little doubt from the authors' comments that the difference in P.R.T. means is almost entirely caused by the early reactions of the anxiety and hysteria cases. They also reported a high degree of test-retest reliability for eleven neurotics tested on 6 successive days.

Preliminary results for two groups of female psychoneurotics matched for age, one consisting exclusively of anxiety cases ($N = 15$), the other of hysterics, neurasthenics and obsessionals ($N = 12$), have been reported by Hemphill, Hall & Crookes (1952). These show that, although both mean P.P.T. and mean P.R.T. are lower in the anxiety cases than in the other neurotics, only the P.R.T. means are statistically significantly different. This, in part, confirms the findings of Chapman *et al.* (1947), and emphasizes the need for a fuller study of the subgroup differences.

Chapman *et al.* (1946) found the same tendency for lower than normal P.R.T.'s in neurotic groups consisting of thirty acute and sixty-two chronic anxiety cases. Although the differences are significant, there is again considerable overlap. Analysing their results qualitatively, the authors found that those patients complaining of general 'nervousness' have the lowest P.R.T., whereas those complaining chiefly of breathlessness, chest pain, weakness, etc., are nearer to the normal.

As a result of their findings, the authors suggested that the test might have some differential diagnostic value. This suggestion was followed up by Schilling & Musser (1948) on seventy-one psychoneurotic patients. Their results only confirm the previous findings of significant difference in P.R.T. means between the neurotics and a group of 138 controls, but again with a large overlap in the distributions. From these findings, it seems probable that the diagnostic value of the test may be not so much in differentiating patients from normals as in aiding in the establishment of one form of syndrome as primary rather than another.

Schilling & Musser (1949*a*) also investigated P.R.T.'s in twenty-two non-psychotic female patients diagnosed as menopausal syndrome and in twenty-three peptic ulcer patients (1949*b*). Both patient groups had significantly lower mean P.R.T.'s than comparable normals. It would seem, therefore, that the lowered P.R.T.'s shown by these various groups of patients are due to some general factor of emotional disturbance. It is possible that the peptic ulcer patients might have a heightened sensitivity for any form of pain as a result of the intermittently painful condition of their illness, although a recent study by Kennard (1952) of patients suffering from severe chronic painful conditions having a genuine organic basis shows that these patients differ only in greater variability of reaction (as indicated by G.S.R.) and not in any consistent hyper-reactivity. Whereas distraction produced by an intense pain in another part of the body has been found to raise the forehead P.P.T. by 35 % (Hardy *et al.*, 1940), there is no adequate evidence as yet on the effect, for instance, of hypochondriacal or other psychogenic pains on the P.P.T. or P.R.T.

The problem of reactivity to pain in psychoneurosis has been considerably advanced in a number of studies from the Allan Memorial Institute of Psychiatry, Montreal. The standard procedure in this work has not involved threshold determination, but the concurrent recording of various levels of psychophysiological response in the patient during a series of twelve thermal stimulations on the forehead, all of which were deemed by the experimenters to be above the perceptual threshold for pain.

The findings from this work are in agreement with that already reviewed, in that psychoneurotics tend consistently to overreact both at a motor and at an autonomic level of response. Malmo, Shagass, Davis, Cleghorn, Graham & Goodman (1948) found that their neurotics showed a significantly greater amount of finger movement during pain stress than paired normal controls, while they also showed a higher frequency of anticipatory responses, as indicated by change in skin resistance. A study by Malmo & Shagass (1949*a*) of two groups of patients, one having cardiovascular complaints, the other having head and neck pains, is of particular interest in demonstrating a relative 'specificity of reaction' as between the groups. The former group showed a significantly greater mean heart rate, mean heart rate variability and median respiratory variability, while the latter showed significantly greater muscle potential scores. This is a finding of some importance in demonstrating how reactivity to pain stress in psychiatric patients with somatic complaints tends to manifest itself in increased responsivity in the related physiological system or mechanism.

Three other studies from the same Institute (Malmo & Shagass, 1949*b*; Malmo, Shagass & Heslam, 1951; Malmo & Shagass, 1952) have shown clearly that psychoneurotics differ from the normal with respect to pain reactivity not only in degree of anticipatory reaction during the stimulus but also very markedly in duration of physiological disturbance. This is shown by the fact that a high blood-pressure level, recorded early in the pain situation, tends to persist, or to increase, at a much later phase of the experiment in the neurotics than in the normal subjects.

While the evidence as to the over-reactivity of psychoneurotics under pain stimulation is convincing and consistent, the findings in most, but not all, of the work that they do not differ from the normal in pain perception is somewhat unexpected. On this point, the actual form of instruction and the general experimental setting, as well as the severity of the neurotic disorder, are probably crucial factors in determining whether or not a neurotic patient anticipates verbally the feeling of pain.

(ii) *Response to pain in schizophrenia and depression*

Very little work has so far been reported on pain response in psychotic conditions, partly, it seems, because it has been assumed by some investigators that it is impossible to obtain thresholds from such patients. In the sense of obtaining accurate psychophysical thresholds, this is obviously true, but a few studies have been carried out in which controlled observations of response to heat-pain stimuli have produced evidence of some psychological interest. Hemphill *et al.* (1952) studied pain perception and reaction in two groups of female depressives. The first group (endogenous depressions) tended to have a very high P.P.T. and P.R.T., while the second group (involutional, agitated, and other psychotic depressions), showed an extremely wide range of response overlapping considerably with some of the female neurotic patients also studied.

Reactivity to pain in schizophrenia has been the subject of investigations at the Allan Memorial Institute, Montreal, where the same pain-stress situation and concurrent physiological recordings were used as have already been described in the work on psychoneurotics. Malmo & Shagass (1949*b*) found that early schizophrenics tended to show a high degree of reactivity to the pain stimuli, coming nearest to the anxiety cases in this respect. On the other hand, Malmo, Shagass & Davis (1951) and Malmo, Shagass & Smith (1951) found that chronic schizophrenics, as compared with normals, neurotics and acute psychotics, showed a generally low responsiveness as measured by mean arm tension change and very rarely pressed the button down, as instructed to do, to signal when pain was felt. At the same time, measures of spontaneous background physiological reactivity were as high for the chronic schizophrenics as for the other groups, and the authors concluded that these patients differed primarily in their very greatly reduced reactivity at a purposive, directed level of response.

These studies of reactivity to pain tend to ignore the perceptual-conceptual aspects of the total stimulus-response situation, the failure to produce an appropriate defence reaction both in the chronic schizophrenic and in some depressive patients being probably due primarily to a discriminative and conceptual failure rather than to a motor retardation. Both the receptor and the effector aspects of pain, and their interrelations, in different clinical states, require much further detailed study.

(iii) *Effects of lobotomy*

The effects of lobotomy upon pain appreciation and reaction to pain have been assessed chiefly for two types of case. In the first type, the operation is carried out in the normal way upon psychiatric patients, their responses to the standard heat-pain test being observed before and at various intervals after operation. In the second group of cases, the operation is carried out for the relief of intractable pain due to actual organic disease, such as cancer. In this group, the evidence of alteration in the patient's response to pain is entirely clinical, except for one threshold investigation, and is concerned with the general observations as to the patients' 'reduced apprehension', 'reduced fear of death', and so on.

The changes in reaction to pain in psychiatric patients have been studied by Chapman, Rose & Solomon (1948). Using the standard heat-pain apparatus and procedure, they obtained motor withdrawal thresholds for twenty-three patients, before and after lobotomy. Their results showed that the patients winced or withdrew from the apparatus at

less intense stimuli after lobotomy than before. Statistically, the differences between pre-operative level and a 15- to 60-day post-operative level were found to be highly significant. Although these differences seemed to persist at the 3- to 6-month, and the 6- to 12-month retests, there was a tendency for the discrepancies to be reduced, reaction thresholds tending to return to the pre-operative level. The authors therefore supposed that their results might represent only a temporary effect of the operation. The post-operative lowering of reaction threshold did not appear to be due to more specific factors, such as sensitization of the forehead skin as a result of the operation, because very similar results were obtained when the stimulus was given on the back of the hand. That this reaction is qualitatively dissimilar to the early, anticipatory kind of reaction shown by some anxiety cases, is shown by Malmo & Shagass (1950) who used their stress tests, including that for pain, on patients before and after lobotomy. Their results gave objective evidence of lowered tension, after operation, during the series of pain stimuli, this being shown in reduced finger movement, reduced disturbance for neck muscle potentials and steadier heart rate.

Chapman, Rose & Solomon (1950) followed up their patients for a further year after the operation. Their results confirmed the trend they had already noted, namely that the motor reaction levels tended to return to the pre-operative level. They suggested that this might be due to some other mechanism in the brain taking over the inhibitory function which seemed to be temporarily reduced in the first months after operation.

The observations made on the second type of case where lobotomy is carried out for relief of intractable pain are chiefly of interest as clinical demonstration of a greatly changed attitude towards the pain of a disease, and possibly towards pain stimulation of any kind. It is suggested further, that it is only the 'attitude' and not the 'sensitivity' of the patient that is altered. Freeman & Watts (1946) reported on five cases where the pains, for which there was considerable structural basis, were severe, constant and disabling. As it was felt that the emotional response to the pain, the fear of pain, had become more disabling than the pain itself, it was hoped that psychosurgery would relieve much of the emotional stress. The authors maintain that the procedure had good effect in that it altered these patients' emotional response to pain without changing materially their ability to feel pain. Similar observations were made by Watts & Freeman (1948) and Falconer (1948). Freeman & Watts (1948) also described a case of involutional depression, with great pain experienced out of all proportion to the physical cause (psychalgia) for whom the operation greatly reduced the suffering. They reported on another post-lobotomy case whose '. . . perception of pain is as keen as ever, but he now speaks of the pains as twinges and can laugh them off' (p. 750). Confirming the anxiety reduction produced by the operation, Dynes & Poppen (1949) reported that their patients, who had shown great fear of death, would, after lobotomy, talk freely about their disorder and approaching death with apparent complete detachment and absence of feeling or concern.

The only actual cutaneous pain threshold investigation on patients suffering from intractable somatic pain seems to have been carried out by King, Clausen & Scarff (1950). Their six patients were all operated on solely for the relief of pain, the operation being that of unilateral prefrontal lobotomy. The patient was instructed to attend to the stimulus during the exposure of the heat on the forearm and to report whatever sensations he experienced during this period. Five of the six patients experienced relief of pain

following operation, and four of these had lowered cutaneous thresholds bilaterally in comparison with pre-operative levels.

This study seems to confirm the findings of Chapman and others (Chapman, Rose & Solomon, 1950; Malmo & Shagass, 1950) with psychiatric patients, the general conclusions being that lobotomy, at least temporarily, lowers the reaction threshold to pain stimuli while at the same time reducing anticipation and the emotional concomitants of the pain sensation. Cobb's (1943) general explanation for the relief of psychotic symptoms through lobotomy would seem to fit the present evidence. He believes that the operation limits the spread of stimuli to different cortical areas, thereby preventing the integration of the present stimuli with past memories and habitual responses. This is similar to Goldstein's (1949) psychological explanation in terms of the more 'concrete' attitude of these patients, which is reflected also in their reduced conceptual ability. Further investigation of the relation of this 'primitive' kind of response to pain to conceptual level might be of interest, while the reactions of deteriorating organic psychotics might be usefully compared.

The chief interest of the work so far carried out on pain response in psychiatric conditions lies in the clear demonstration of the important function of the emotional and conceptual setting of the pain stimulus. At one extreme, the emotional disturbance of the neurotic may produce a lowering of reaction threshold to pain, and a physiological over-reactivity that is long-lasting and intense. At the other extreme, the apathy and inertia of some schizophrenics and depressives leads to high threshold and under-reactivity, while the lobotomy results emphasize how relatively indifferent even a previously extreme pain can seem when it is operatively dissociated from its emotional context.

IV. Conclusions

It will be apparent from this survey that the psychological aspects of pain, both in normal and in psychiatric conditions, have been thoroughly investigated in so far as they derive directly from the standard psychophysical procedures such as are used on other sensory threshold problems. The relevance of certain general factors, such as age, sex, verbal intelligence level, has been indicated in a number of experiments, as also has the importance of 'attitude', whether induced by experimental variation or subjectively derived from a standard situation. There is at present, however, little experimental evidence as to the precise nature of the effects of these factors upon response to pain, most of the work in which they are mentioned not being designed or statistically treated so as to bring out their relative significance. Similarly, the relation of the attitudes postulated to other personality characteristics, whether primarily learned or constitutionally derived, remains unformulated. The initial general hypothesis would seem justifiably to be that the major causes of variation in pain response are due to central factors—differences in capacity to inhibit response or to perceive the whole stimulus situation in conceptual setting.

In psychiatric conditions, the evidence for over-reactivity to pain stimulation in neurotics and for under-reactivity in some schizophrenics and depressive states accords with clinical expectation. In such conditions, however, the perceptual-conceptual aspects of the pain situation are of at least equal importance, and studies of these and the interrelationships between various measures of pain response and other forms of responsiveness, such as those of reaction-time, should help to establish more clearly the nature of the central processes involved. Studies of lobotomized patients have shown how altered the response to pain can be, at least temporarily, when the so-called 'pain'

stimulation is isolated from its emotional context, and it is possible that experimental isolation of pain stimuli, by variation of surrounding instructional and objective conditions, may contribute to the understanding of the emotional factors intervening between pain stimulus and pain response.

References

Bigelow, N., Harrison, I., Goodell, H. & Wolff, H. G. (1945). Studies on pain: quantitative measurements of two pain sensations of the skin. *J. Clin. Invest.* xxiv, 503–12.

Chapman, W. P. (1944). Pain sensitivity in normal control subjects and psychoneurotics. *Psychosom. Med.* vi, 252–7.

Chapman, W. P., Cohen, M. E. & Cobb, S. (1946). Pain in neurocirculatory asthenia, combat fatigue and anxiety neurosis. *J. Clin. Invest.* xxv, 890–6.

Chapman, W. P., Finesinger, J. E., Jones, C. M. & Cobb, S. (1947). Measurements of pain sensitivity in patients with psychoneurosis. *Arch. Neurol. Psychiat.* lvii, 321–31.

Chapman, W. P. & Jones, C. M. (1944). Variations in cutaneous and visceral pain sensitivity in normal subjects. *J. Clin. Invest.* xxiii, 81–91.

Chapman, W. P., Rose, A. S. & Solomon, H. C. (1948). Measurements of heat stimulus producing motor withdrawal reaction in patients following lobotomy. *Res. Publ. Ass. Nerv. Ment. Dis.* xxvii, 754–68.

Chapman, W. P., Rose, A. S. & Solomon, H. C. (1950). A follow-up study of motor withdrawal reaction to heat discomfort in patients before and after lobotomy. *Amer. J. Psychiat.* cvii, 221–4.

Chapman, W. P. & Smith, L. H. (1948). Effects of hypoxia and hypercapnia on perception of cutaneous pain. *J. Clin. Invest.* xxvii, 299–304.

Child, I. L. (1950). Relation of somatotype to self-ratings on Sheldon's temperamental traits. *J. Personality*, xviii, 440–53.

Clausen, J. & King, H. E. (1950). Determination of pain threshold on untrained subjects. *J. Psychol.* xxx, 299–306.

Cobb, S. (1943). Borderlands of Psychiatry. *Harvard University Monograph*, no. 4.

Dunbar, F. (1947). See Chapman, Finesinger, Jones & Cobb.

Dynes, J. B. & Poppen, J. L. (1949). Lobotomy for intractable pain. *J. Amer. Med. Ass.* cxl, 15–18.

Edwards, W. (1950). Recent research on pain perception. *Psychol. Bull.* xlvii, 449–74.

Falconer, M. A. (1948). Relief of intractable pain of organic origin by frontal lobotomy. *Res. Publ. Ass. Nerv. Ment. Dis.* xxvii, 706–14.

Freeman, W. & Watts, J. W. (1946). Pain of organic disease relieved by prefrontal lobotomy. *Lancet*, 29 June, i, 953–5.

Freeman, W. & Watts, J. W. (1948). Pain mechanisms and the frontal lobes: a study of prefrontal lobotomy for intractable pain. *Ann. Intern. Med.* xxviii, 747–54.

Furer, M. & Hardy, J. D. (1950). The reaction to pain as determined by the G.S.R. *Res. Publ. Ass. Nerv. Ment. Dis.* xxix, 72–89.

Goldstein, K. (1949). Frontal lobotomy and impairment of abstract attitude. *J. Nerv. Ment. Dis.* cx, 93–111.

Hall, K. R. L. & Stride, E. (1953). The varying response to pain in psychiatric disorders. *Brit. J. Med. Psychol.* (To be published.)

Hardy, J. D., Goodell, H. & Wolff, H. G. (1951). Influence of skin temperature on pain threshold. *Science*, cxiv, 149–50.

Hardy, J. D. & Javert, C. T. (1949). Studies on pain: measurements of pain intensity in childbirth. *J. Clin. Invest.* xxviii, 153–62.

Hardy, J. D., Wolff, H. G. & Goodell, H. (1940). Studies on pain: a new method for measuring pain threshold. *J. Clin. Invest.* xix, 649–57.

Hardy, J. D., Wolff, H. G. & Goodell, H. (1943). Relation of attitude and suggestion to the perception of and reaction to pain. *Res. Publ. Ass. Nerv. Ment. Dis.* xxiii, 434–48.

Hardy, J. D., Wolff, H. G. & Goodell, H. (1947). Studies on pain: discrimination of differences in intensity as basis of a scale of pain intensity. *J. Clin. Invest.* xxvi, 1152–8.

Hardy, J. D., Wolff, H. G. & Goodell, H. (1948). Studies on pain: some quantitative aspects of the dol scale. *J. Clin. Invest.* xxvii, 380–6.

Hardy, J. D., Wolff, H. G. & Goodell, H. (1950). Experimental evidence on the nature of cutaneous hyperalgesia. *J. Clin. Invest.* xxix, 115–40.

Hemphill, R. E., Hall, K. R. L. & Crookes, T. G. (1952). Preliminary report on fatigue and pain tolerance in depressive and psychoneurotic patients. *J. Ment. Sci.* xcviii, 433–40.

Herget, C. M., Granath, L. P. & Hardy, J. D. (1941). Thermal sensation and discrimination in relation to intensity of stimulation. *Amer. J. Physiol.* cxxxiv, 645–55.

Hines, E. A. & Brown, G. E. (1932). A standard stimulus for measuring vasomotor reactions. *Proc. Mayo Clin.* vii, 332. (Cited by Wolf, S. & Hardy, J. D.)

Ischlondsky, N. E. (1949). *Brain and Behaviour.* London: Kimpton.

Janoff, I. Z., Beck, L. H. & Child, I. L. (1950). Relation of somatotype to reaction time, resistance to pain, and expressive movement. *J. Personality,* xviii, 454–60.

Jenkins, W. L. (1951). Chapter on 'Somesthesis'. In *Handbook of Experimental Psychology,* ed. S. S. Stevens. New York: Wiley.

Kennard, M. A. (1952). Responses to painful stimuli of patients with severe chronic painful conditions. *J. Clin. Invest.* xxxi, 245–52.

King, H. E., Clausen, J. & Scarff, J. E. (1950). Cutaneous thresholds for pain before and after unilateral prefrontal lobotomy. *J. Nerv. Ment. Dis.* cxii, 93–6.

Lanier, L. H. (1943). Variability in the pain threshold. *Science,* xcvii, 49–50.

Lewis, T. (1942). *Pain.* New York: Macmillan.

Livingston, W. K. (1943). *Pain Mechanisms.* New York: Macmillan.

Malmo, R. B. & Shagass, C. (1949*a*). Physiologic study of symptom mechanisms in psychiatric patients under stress. *Psychosom. Med.* xi, 9–24.

Malmo, R. B. & Shagass, C. (1949*b*). Reaction to stress in anxiety and early schizophrenia. *Psychosom. Med.* xi, 25–9.

Malmo, R. B. & Shagass, C. (1950). Behavioural and psychological changes under stress after operations on the frontal lobes. *Arch. Neurol. Psychiat.* lxiii, 113–24.

Malmo, R. B. & Shagass, C. (1952). Studies of blood pressure in psychiatric patients under stress. *Psychosom. Med.* xiv, 82–93.

Malmo, R. B., Shagass, C., Davis, J. F., Cleghorn, R. A., Graham, B. F. & Goodman, A. J. (1948). Standardized pain stimulation as controlled stress in physiologic studies of psychoneurosis. *Science,* cviii, 509–11.

Malmo, R. B., Shagass, C. & Davis, J. F. (1951). Electromyographic studies of muscular tension in psychiatric patients under stress. *J. Clin. Psychopath.* xii, 45–66.

Malmo, R. B., Shagass, C. & Heslam, R. M. (1951). Blood pressure response to repeated brief stress in psychoneurosis: a study of adaptation. *Canad. J. Psychol.* v, 167–79.

Malmo, R. B., Shagass, C. & Smith, A. A. (1951). Responsiveness in chronic schizophrenia. *J. Personality,* xix, 359–75.

Morgan, C. T. & Stellar, E. (1950). *Physiological Psychology,* 2nd ed. New York: McGraw-Hill.

Rey, A. (1943). Recherches sur la sensibilité cutanée normale et pathologique. *Arch. Psychol., Genève,* xxx, 1–63.

Schilling, R. F. & Musser, M. J. (1948). Pain reaction thresholds in psychoneurotic patients. *Amer. J. Med. Sci.* ccxv, 195–7.

Schilling, R. F. & Musser, M. J. (1949*a*). Pain reaction thresholds in the menopausal syndrome. *Amer. J. Med. Sci.* ccxviii, 204–6.

Schilling, R. F. & Musser, M. J. (1949*b*). Pain reaction thresholds in patients with peptic ulcer. *Amer. J. Med. Sci.* ccxviii, 207–8.

Schumacher, G., Goodell, H., Hardy, J. D. & Wolff, H. G. (1940). Uniformity of the pain threshold in man. *Science,* xcii, 110–12.

Sheldon, W. H. & Stevens, S. S. (1942). *The Varieties of Temperament.* New York: Harper.

Stone, L. J. & Jenkins, W. L. (1940). Recent research in cutaneous sensitivity: I. Pain and temperature. *Psychol. Bull.* xxxvii, 285–311.

Watts, J. W. & Freeman, W. (1948). Frontal lobotomy in the treatment of unbearable pain. *Res. Publ. Ass. Nerv. Ment. Dis.* xxvii, 715–22.

Wertheimer, M. & Ward, W. D. (1952). Influence of skin temperature: a confirmation. *Science,* cxv, 499.

West, L. J., Neill, K. C. & Hardy, J. D. (1952). Effects of hypnotic suggestion on pain perception and G.S.R. *Arch. Neurol. Psychiat.* lxvii, 694–5.

Wolf, S. & Hardy, J. D. (1941). Studies on pain: observations on pain due to local cooling. *J. Clin. Invest.* xx, 521–33.
Wolff, H. G. & Goodell, H. (1943). Pain threshold in man. *Res. Publ. Ass. Nerv. Ment. Dis.* xxiii, 1–15.
Wolff, H. G. & Hardy, J. D. (1947). On the nature of pain. *Physiol. Rev.* xxvii, 167–99.
Wolff, H. G., Hardy, J. D. & Goodell, H. (1940). Measurements of the effect of morphine, codeine, and other opiates on the pain threshold. *J. Clin. Invest.* xix, 659–80.
Wolff, H. G. & Wolf, S. (1948). *Pain.* Springfield: Thomas.

(*Manuscript received* 17 *October* 1952)

NOISE, PACED PERFORMANCE AND VIGILANCE TASKS

By D. E. BROADBENT

Applied Psychology Research Unit, Cambridge

I. *Introduction* (p. 295). II. *Procedure* (pp. 295–296). III. *Results* (pp. 296–299). IV. *A comprehensive account of vigilance tasks* (pp. 299–302). V. *Summary* (pp. 302–303). VI. *References* (p. 303).

I. Introduction

Most previous workers on noise have not succeeded in establishing any serious detrimental effect on performance (Berrien, 1946; Kryter, 1950). It was found recently, however, that marked effects of noise could be demonstrated provided that the subjects used were not selected by academic examination or similar means, and that the same task was continued in noise for a prolonged period (Broadbent, 1953). These points had not been covered adequately by previous work. In addition, the tasks performed by the subjects in the recent research were vigilance tasks, that is, they required visual watch-keeping for occasional signals, and such tasks are somewhat different from the functions previously tested. Clearly this effect requires further investigation: more particularly since an easier task of watching for warning lights (the Lights Test) only showed an effect on some individuals, while a harder task of watching a complex dial display (the Dials Test) showed an effect on all subjects.

Vigilance tasks have been used in several situations in the past few years, and there are some discrepancies in the findings which also need clarifying. The Clock Test of Mackworth (1950) showed a rapid 'fatigue' decrement when a subject watches for signals from a single source such as a radar set. The Dials Test of Broadbent (1950), however, did not show such a decrement in the first few runs when a subject watches for signals from many sources such as the dials arranged round a control room. As rapid perceptual 'fatigue' is of extreme importance both in practice and in theoretical accounts of learning and other phenomena, it was highly desirable that the conditions under which it appears should be known more exactly.

The present paper describes some experiments on prolonged performance in quiet and in noise, on a task which is not a vigilance one. The data from these experiments then make possible a rational account of previous vigilance experiments and of the differences between them, and also of the noise effect which previous workers have missed.

II. Procedure

The task employed in all the following experiments demanded a 5-choice serial reaction. Five neon lamps were displayed to the subject in the form of an equilateral pentagon, and he had before him five contacts arranged in the same way which could be touched with a stylus held in the hand. The task at any instant was simply to touch the contact corresponding to the one lamp which was alight. In one condition (the 'paced' one) the lamp went out and was replaced by another once every second. In the other condition (the 'unpaced'), this change occurred as soon as the subject touched one of the contacts: if the choice was incorrect the light still changed, so that both conditions were alike in the

amount of knowledge of results present. The next light to appear was chosen by a uniselector mechanism designed and built by Dr A. Leonard.

The apparatus used for the preliminary experiments carried the display at eye-level facing the subject while the contacts lay horizontally on a table in front of him. In these first groups, all subjects were told that the task was similar to that required of certain radar operators, and were told to keep on working as accurately as possible until they could do so no longer. (In fact, none of them stopped completely in periods of about an hour and a half, and the experimenter terminated the session at this approximate time.) After they had practised until they were working smoothly and accurately, they were then left alone without watch or cigarettes and the experimenter observed their work from the recording apparatus in the other room. Group I (nine subjects) followed this routine with the unpaced condition, and group II (six subjects) with the paced.

In the main experiment, each lamp was placed just beside its corresponding contact and the combined panel mounted at an angle in front of the subject. Group III (eighteen subjects) performed two half-hour periods of work each on this task, being told as before that the task was one demanded of radar operators, but in this case being asked to work as fast and as accurately as possible for half an hour. The unpaced condition was used throughout. Half the subjects carried out their first run in quiet and their second in noise, and the other half reversed this order. Quiet here means 70 db. and noise 100 db. re 0·0002 dyne/cm.^{2}, the details of the production and spectrum of the sounds being given by Broadbent (1953). Roughly speaking, the noise contained equal energy in all frequency bands from 100 to 5000 c.p.s.; it remained steady throughout the test, and had no directional characteristics.

Group IV (fourteen subjects) were treated as group III, but each was seen individually by the writer before the experiment, and informed with circumstantial detail and prepared graphs that it was known that noise improved both accuracy and speed on other similar tasks, but that it was important for the design of radar sets to know exactly how large this improvement was on this type of equipment.

All the subjects were Naval ratings between the ages of 18 and 30; in the main group two were stokers, four seamen, and twelve Fleet Air Arm ratings. All were therefore accustomed to noise.

III. Results

(*a*) *Pacing and prolonged performance.* Group I, the unpaced group, showed no significant change in the output of correct responses from that during the first 10 min., until the sixth 10 min. was reached. This sixth 10 min. was just significantly worse than the first, $t=2{\cdot}32$, while the 0·05 level is 2·31. It was noticeable, however, that short pauses of 2 sec. or more occurred at intervals in the performance, and the incidence of such pauses increased markedly much earlier in the period. In the first 10 min. there was an average of 5 such pauses, and in the second of $8\frac{1}{2}$: this increase is easily significant, $t=2{\cdot}84$ after normalizing the distribution by taking log $(1+x)$, where x is the number of pauses in a period for one subject. The 0·05 level is 2·31. Only one subject showed a change in the opposite direction.

Group II, which was paced, showed a fall-off in the output of correct responses after the first 10 min. The average fell from 534 in the first 10 min. to 493 in the second, $t=2{\cdot}841$, while the 0·05 level is 2·57. Every subject showed this deterioration. This

result would be expected from that in group I. In the unpaced condition momentary pauses in the work are there shown to increase after 10 min., but the subjects are able to compensate for them by increasing their rate of work between pauses. In the paced condition, however, a momentary pause causes a loss of the opportunity to make a correct response, and there is no possibility of compensating by an increased rate at other times.

The effect of prolonged performance in increasing gaps in a continuous task has of course been reported previously (Bills, 1931). In the present case a fixed time limit was used to define a gap or 'block', rather than a criterion varying with the individual subject, but the results would be expected to be parallel. There is here an apparent conflict, however, with the work of Bills & Shapin (1936): these workers reported that there was no difference in output decrement between paced and unpaced conditions, and that the latter gave a larger increase in 'blocks'. The explanation of this discrepancy is that Bills & Shapin used a rate of pacing too fast for the subjects to keep up with continuously, while the present experiment used the same average rate as that found for the unpaced condition. In the previous work, therefore, subjects could compensate for pauses by spurting between them as in the unpaced condition of the present experiment. Bills & Shapin also point out that they gave the paced condition second to all subjects, despite the marked practice effects.

Table 1. *Indices of effect of prolonged work (all eighteen subjects of main group)*

	Run I		Run II	
	1st 10 min.	2nd 10 min.	1st 10 min.	2nd 10 min.
Output of correct responses	15,308	15,455	17,114	16,641
Gaps in responding	100	151	50	67
Incorrect responses	235	311	280	543

It remains to note that a 'fatigue' effect on gaps but not on output was also found in the main group as far as the first run was concerned. Only one subject of the eighteen showed a decrease in the number of short gaps after the first 10 min., and fifteen showed an increase. The distribution of differences is somewhat skew, but this result is, of course, significant at the 0·01 level by χ^2, using only information on the direction of the differences. In the second run, however, there appeared to be no uniform relation between the number of pauses and time since the beginning of the run. As will be pointed out in the next section, this is not surprising in view of recent work on 'anticipation' (Poulton, 1950). In neither run was there a significant change in either direction in the output of correct responses. The detailed results are given in Table 1.

(*b*) *Errors and prolonged performance.* In work on 'blocks', errors have usually been regarded as associated with 'blocks'. There was an increase in errors in group I from the first 10 min. to the second, with only one subject showing an opposite effect; but his results conflicted so markedly with those of the remainder that the increase was not significant, $t = 1{\cdot}33$ instead of 2·31 at the 0·05 level. The increase continued to appear in the main group in both runs; when the results from the two runs are pooled, sixteen subjects made more errors in the second 10 min. than in the first, and two reversed this trend. This result again is highly significant by χ^2, reaching the 0·01 level, without considering the distorted distribution of the differences. The average number of errors in each run was 14·5 for the first 10 min. and 20·5 for the second 10 min., each out of about 350 correct responses.

The fact that errors continue to act as an index of 'fatigue' decrement in the second run, while gaps do not, is also not unexpected in the light of recent work on anticipation (Poulton, 1950). A recorded error in this task is not a failure of co-ordination or similar breakdown of response, but the perfect execution of a response which has been chosen incorrectly or without reference to the stimulus situation. Now in familiar tasks an individual frequently initiates responses before the stimulus requiring them has appeared, and this function enables him to continue responding continuously even when incoming sensory information is temporarily withheld (loc. cit.). It has been common to regard 'blocks' as failures of attention, and if this is the case, it would be expected that as a task becomes familiar the subject would carry on working through such failures by the same function as that used in anticipation. Since the stimulus changes are in this case objectively unpredictable, this would result in errors: and so in a practised task errors rather than gaps would show effects of prolonged performance. The relation between paced and unpaced conditions would of course be unaffected by this change.

(*c*) *The effect of noise.* Errors, which continued to show an effect of prolonged performance in the second run, were markedly more frequent in noise. The first 25 min. of each run were used for scoring, as the last 5 min. had to be cut short for extraneous reasons on some subjects: in this 25 min. the average number of errors per subject was

Table 2. *Numbers of subjects making more errors in noise and in quiet*

	Many less in run 2	Few less in run 2	Few less in run 1	Many less in run 1
Subjects having first run in quiet	0	1	2	6
Subjects having first run in noise	2	3	4	0

37 in quiet and 57 in noise. The distributions were even more skew than those met previously, so a coarse classification was made. A 'minor' difference was defined as one of six errors or less, and a 'major' one as seven errors or more, and Table 2 was then compiled. This table is highly significant using τ, S-Continuity Correction of $S = 2{\cdot}75$ as opposed to 2·58 for the 0·01 level. It will be noted that all the major differences favour quiet: in the extreme case one individual made five times as many errors in noise as in quiet.

A point of interest is that the increase in errors in noise does not appear early in the run: Fig. 1 shows that the number of subjects favouring quiet and favouring noise in the first 5 min. might actually have been interpreted as a superiority for noise, a point of importance in view of the short tests used in some of the better-controlled noise investigations (Kryter, 1950). Similar results were found when vigilance tasks were used (Broadbent, 1953).

The output of correct responses showed no noticeable effect of noise.

Group IV, the suggestion group, also showed an increase in errors in noise, the means being 24 for quiet and 35 for noise. This difference is smaller than that in group II, but of much the same size relative to the number of errors in quiet. By a test of similar type to that used in the main group, the difference is significant at the 0·05 level assuming that only a difference in this direction would be accepted. The significance is improved if we consider only the worst 5 min. in any run, when the noise run is found to reach a lower level, S-Continuity Correction/S.E. of $S = 2{\cdot}01$ instead of 1·96 at the 0·05 level without any assumption as to the direction of the difference.

As was noted in group III, the effect of noise is not uniform throughout the work-period, and it may be that a suggestion favouring noise increases the length of time for which no effect appears. But the slight difference found may easily be due simply to sampling differences between the groups of subjects.

(*d*) *The Triple Tester Index.* This index is obtained by giving a subject nine runs on the Triple Tester (a tracking task), informing him after each run of his score and asking him to estimate his future score. The difference between the estimate and the immediately previous attainment is then summed regardless of sign for all estimates except the very first; and the resulting score has been shown to correlate negatively with the amount of decrement found with prolonged performance in a vigilance task such as the Twenty Dials or Twenty Lights. It is also related to effects of noise (Broadbent, 1953).

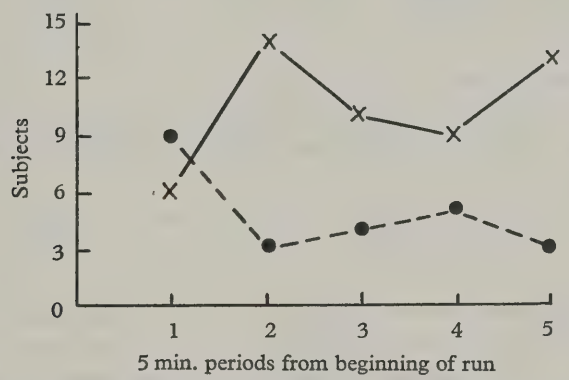


Fig. 1. The effect of time since onset. Number of subjects working more efficiently under each condition in successive 5 min. from the beginning of the run. ●, subjects making less mistakes in noise; ×, subjects making less mistakes in quiet.

In the present case this index was obtained for each subject in group III, and showed a negative correlation with amount of decrement in output of correct responses between the first 10 min. period and the second. If we sum the decrement for both runs for each subject, calculate τ for each of the two subgroups and pool the values, we obtain a τ of $-0{\cdot}44$, S/S.E. of $S = 2{\cdot}36$, while the 0·05 level is 1·96. The individual differences in effect of prolonged performance appear therefore to be similar in this continuous task to those on vigilance tasks.

IV. A comprehensive account of vigilance tasks

As has been said, it has been common to explain 'blocks' by supposing that attention cannot be maintained steadily on one object for a long period: as time goes by it must change its object and may then return to the task. On a different level of explanation, concepts of physiological fatigue have been employed, and the two views are by no means inconsistent with one another (Bills, 1943). Taking the former view, we may draw an analogy with vigilance tasks: defining attention for those dissatisfied with a subjective concept, as the selection for response of one part of the total stimulus situation. It has been pointed out by Hick (in an unpublished thesis) that the amount of information furnished by the sense organs is greater than that which can be handled by an analyser the size of the human brain; so that such a process of selection is necessary.

The Clock Test delivers once a second a stimulus to which the subject must decide 'No response' or 'Response' (Mackworth, 1950). If his attention is not directed to the task for a short pause of 2 sec. or so, the stimulus has gone and he cannot catch it up. This is comparable to the 'paced' condition in continuous performance, and we would therefore expect a decrement in the number of correct responses as the task proceeds. This in fact occurs. When signals are not transitory, however, but come from a number of separate sources and remain on view until a response is made, the situation becomes less and less paced; and this is the case in the Dials and Lights Tests (Broadbent, 1950, 1953). Here the subject scans the various sources of signals, making a similar series of decisions to respond or not to do so, but he scans at his own rate. A *short* failure of attention may be compensated by a spurt in the rate—the extent to which such compensation is possible will be determined by the difficulty of each separate decision. We would expect, therefore, that no decrement with prolonged watch would appear on the Dials and Lights Tests, which is the case.

There are two discrepancies in this analogy:

(*a*) In the 5-choice reaction all stimuli are equally like those before and after them, while in the vigilance tasks most stimuli are of one kind and only an occasional one (the signal) of another.

(*b*) The times involved are shorter in the 5-choice case than in the Dials. The latter still shows no decrement after $1\frac{1}{2}$ hr., whilst the unpaced 5-choice task does begin to show one after an hour.

One of these discrepancies, however, explains the other. Our original principle was that drawn from experiments on 'blocks': that an unchanging situation will not continue to evoke a continuous series of responses indefinitely. There is some further evidence for this from the finding on the Dials that, as a period of monotonous watch continues, the relative priorities of different parts of the display change (Broadbent, 1950, p. 8). If this is so, then discontinuity in a task is likely to reduce the number of blocks. The occurrence of an unusual stimulus, the signal, should then improve performance, and it will be recalled that the occurrence of a signal on the Dials partly reinstated the type of performance appearing early in the watch (loc. cit., p. 9). Even when short pauses begin to occur, the Dials subject will still be kept aware of this discontinuity, since the signal will wait for him. It would be expected, then, that the Dials performance would be maintained longer than that on the 5-choice task.

The effect of signals would not be so great on the Clock. On the one hand, the discontinuity would be much less, since signals do not involve an upheaval such as getting up and moving about the room. On the other hand, when signals began to be missed, discontinuity would cease since the subject would never become aware of their occurrence. It would be expected, however, that if the occurrence of signals was emphasized by, say, a succeeding auditory stimulus, and if even missed signals were pointed out to the subject a little later, then the Clock Test would not show so great a decrement in performance, and this is the case (Mackworth, 1950, p. 25). An auditory stimulus not connected with any signal would also have the same effect (loc. cit., p. 22). A short break in performance would equally prove an adequate discontinuity (loc. cit., p. 18).

Subjects who miss fewer signals will show less decrement, and signals which are easier to distinguish and so less often missed will show less decrement (Mackworth, 1950, p. 23; Fraser, 1952). Finally, the presence in the room of a source of continuously changing

stimulation of low intensity but appreciable subjective importance (the experimenter) will reduce the amount of decrement (Fraser, 1953).

Repeated periods of watch will show effects which may be expected to be complicated by individual differences, since those individuals who make an effort to maintain performance on the task may not recover so easily after a break. An individual difference of this type has in fact been shown (Broadbent, 1950, p. 5). Apart from such differences, it is known that the occurrence of signals does not reinstate beginning of watch performance completely (loc. cit., p. 9). Consequently, the effects of continuous performance are cumulative throughout a watch and might be expected to have an effect upon performance on successive days. If so, blocks would tend to appear earlier in the period of watch, together with compensating spurts where these are possible. An unpaced task such as the Dials would then show an increase in the amount of 'fatigue' decrement in each successive watch, and this is the case (loc. cit., App. iv, p. 3). A paced task, such as the Clock, on the other hand, will show a decrease, and this is the case (Carpenter, 1946).

This applies only when the repeated watch is under the same conditions. A change in conditions between runs will be a discontinuity similar to those discussed in the previous paragraph and may cause the second of two runs under different conditions to be better than the second of two under the same condition (Mackworth, 1950, p. 28f.), and to show different 'fatigue' changes.

The foregoing account includes all the major features of vigilance tasks performed in quiet, and derives them rationally from the one principle that response to an unchanging part of the surroundings cannot be maintained continuously for a prolonged period, but will begin to show an uneven distribution in time as the period of work lengthens. If this view is supported by future work, it may be of use both practically and theoretically by describing the conditions under which decrement from prolonged performance will and will not appear. Two final items of evidence may be adduced in support; first, that when successive judgements are made of the time at which a moving object reaches the middle of a projected display, the scatter of judgements by each individual increases very markedly as time goes by (Fraser, 1952): indicating an unevenness of attention in time. Secondly, if a Clock-type task is given using signals which remain visible until the next (non-signal) stimulus appears, reduction of this time will change the task from one which shows no 'fatigue decrement', to one which does (loc. cit.); and this would of course be expected from the change from an unpaced to a paced situation.

As far as noise is concerned, it will be realized that the presence of an intense competing stimulus during a period of continuous performance is likely to increase the frequency of failures in attention to the task portion of the surroundings. The effect of noise will be, in fact, to increase the effect of prolonged performance. By observing for a prolonged period a task in which no compensation is possible for gaps in attention (such as errors in the 5-choice reaction) we will detect this effect; by observing similarly a task in which compensation is difficult (the Dials) we will detect it, though less easily, and by observing a task in which compensation is easy (the Lights) we will detect it only in individuals who do not compensate. There is some evidence in the results from the Lights that individuals who do not deteriorate in noise are being affected but are compensating (Broadbent, 1953). Compensation, together with the possibility of predicting the time of occurrence of the most important signals, will effectively prevent effects of noise from appearing on most tracking tasks, intelligence tests, and so on (Poulton, 1950).

It is, of course, certain that other aspects of the effects of noise remain to be examined. On the one hand, the different reactions of individuals to the increase in failures of attention will furnish material of considerable interest (Davis, 1948; Ryan, Cottrell & Bitterman, 1950). On the other hand, the present work has dealt with tasks which require selective perception of environmental stimuli and shown that this is not always compatible with noise as another part of the environment. It is possible, however, that other forms of behaviour, such as speech, which are determined from within the individual, involve functions which are related to those used in perception (Broadbent, 1952). There may therefore be other and different effects of noise; but that displayed in these results is sufficient to contradict the opinion that effects of noise appear only in popular mythology, and to show why previous investigations have not revealed them.

V. Summary

1. A 5-choice reaction task was set to subjects under the same conditions as a vigilance task, that is with the subject working alone in a monotonous environment without cigarettes or watch. *In preliminary experiments in quiet*, subjects were given no indication of the time for which they would be kept in the room, but were told to continue as long as possible. One group of nine subjects worked under these conditions 'unpaced', that is, with a fresh stimulus being presented each time a reaction was made so that the rate of response depended on the subject. Another group, of six subjects, worked 'paced', that is, with a fresh stimulus being presented automatically every second, whatever the subject's reaction. This rate was about the average 'unpaced' rate.

2. It was found that in the paced condition the output of correct responses fell off after 10 min. work, while in the unpaced condition there was no decrement in output till nearly an hour had gone by. The unpaced subjects, however, showed a marked increase after 10 min. in the number of short pauses in responding, 'short' being here 2 sec. or more. It therefore appeared likely that all subjects were suffering such momentary lapses, but that the unpaced ones could compensate for them by faster work between lapses, while the paced ones could not.

3. *In the main experiment*, eighteen subjects worked at the task unpaced, for half an hour on each of 2 successive days. One day the room was quiet, while on the other day it contained a steady noise of high intensity. The previous results were confirmed for the first day under either condition but pauses ceased to show a decrement in performance with time on the second day: wrong reactions, however, continued to show one. Such errors were far more frequent in noise than in quiet, the average increase being more than half of the number occurring in quiet. A further group of fourteen subjects showed a similar effect of noise despite a preceding suggestion that contrary results were to be expected.

4. On the main group of eighteen subjects the Triple Tester Index was obtained, which has been found to be related to decrement with time in visual watch-keeping. A similar fairly close relation appeared between the index and decrement with time in the output of correct responses.

5. It is suggested that these results allow us to give an adequate interpretation of the peculiarities of vigilance tasks, of differences found between them in the effects of noise and continued exercise, and of the reasons for the apparent failure of many other laboratory tasks to show such effects.

This work was carried out while the author was receiving a grant from the Medical Research Council, and thanks are due to the Royal Navy both for supplying subjects and for providing the very extensive technical facilities required. The influence of Prof. Sir Frederic Bartlett, F.R.S., and of Dr N. H. Mackworth will be readily apparent throughout the paper.

References

Berrien, F. K. (1946). The effects of noise. *Psychol. Bull.* xliii, 141–61.

Bills, A. G. (1931). Blocking: a new principle in mental fatigue. *Amer. J. Psychol.* xliii, 230–45.

Bills, A. G. (1943). *The Psychology of Efficiency.* New York: Harper Bros.

Bills, A. G. & Shapin, M. J. (1936). Mental fatigue under automatically controlled rates of work. *J. Gen. Psychol.* xv, 335–46.

Broadbent, D. E. (1950). The Twenty Dials Test under quiet conditions. A.P.U. Report,[1] No. 130/51.

Broadbent, D. E. (1952). Speaking and listening simultaneously. *J. Exp. Psychol.* xliii, 267–73.

Broadbent, D. E. (1953). Some Effects of Noise on Visual Performance. *Quart. J. Exp. Psychol.*, in the Press.

Carpenter, A. (1946). An experiment to determine whether performance on the Clock Test improves with practice. A.P.U. Report, No. 41/46.

Davis, D. R. (1948). *Pilot Error.* Air Publication 3139*a*. London: H.M.S.O.

Fraser, D. C. (1952). Vigilance Stress. A.P.U. Report, No. 174/52.

Fraser, D. C. (1953). Relationship of an environmental variable to performance in a prolonged visual task. *Quart. J. Exp. Psychol.*, v, 31–2.

Kryter, K. D. (1950). The effects of noise on man. *J. Speech H. Dis.* (*Monograph Suppl.*), No. 1.

Mackworth, N. H. (1950). Researches on the measurement of human performance. *Spec. Rep. Ser. Med. Res. Coun., Lond.*, No. 268.

Poulton, E. C. (1950). Anticipation in open and closed sensorimotor skills. A.P.U. Report, No. 138/50.

Ryan, T. A., Cottrell, C. L. & Bitterman, M. E. (1950). Muscular tension as an index of effort: the effect of glare and other disturbances in visual work. *Amer. J. Psychol.* lxiii, 317–41.

[1] Copies of A.P.U. reports are available from the Medical Research Council and the author.

(*Manuscript received* 19 *August* 1952)

A NEW TYPE OF BEHAVIOUR THEORY

By J. A. DEUTSCH

Institute of Experimental Psychology, University of Oxford

I. *Introduction* (pp. 304–305). II. *The mechanism of performance* (pp. 305–307). III. *Behavioural sequences, innate and learned* (pp. 307–309). IV. *Learning* (pp. 309–311). V. *Quantitative consideration* (pp. 311–313). VI. *Analysers and output* (pp. 313–314). VII. *Conflict and repression* (pp. 314–316). VIII. *Summary* (p. 316). *References* (pp. 317–318).

I. Introduction

The theory put forward in this paper is logically unlike conventional psychological theories. These at best give us equations expressing certain correlations of the phenomena of behaviour. A student of learning theory, for instance, will find much information about the conditions of learning. He will be disappointed if he asks existing theory to explain the phenomena with which it is concerned, for it is merely a generalization about them (e.g. Hull, 1951). There is at present no hypothesis concerning behaviour which is like those ideally employed in other sciences.

Such an explanatory theory is a formal system consisting of elements which are defined by the relations existing between them. This system is furthermore given an identification such that from the system thus identified now follow consequences in terms of the phenomena to be explained. The identification need not be unique. The same formal system may, by making the suitable identifications, be used to explain, say, the behaviour of a machine and the behaviour of an organism.

The formal system or its identification cannot be deduced from the generalizations concerning the phenomena. On the other hand, an acceptable explanatory hypothesis must entail the correct generalizations; it must make the right predictions. One may, of course, predict on the basis of empirical generalizations, but to work on this level is not normally considered ideal in science.

If these points concerning the nature of explanation are correct, then the view that explanation in psychology waits upon neurological advance is wrong. It also follows that a psychological theory need not be physiological speculation. We do not need to give the elements in our system an identification in terms of actual or imagined neural properties or units. This of course is not to deny that the final verification of an explanatory hypothesis in psychology will not come from its direct identification in terms of neuronal systems, but to assert that the framing of a hypothesis can precede its confirmation. We can say that the organization is such-and-such a system without being obliged to state the precise embodiment of this system. It also follows from these points that a theory may, and must, use elements not wholly definable in terms of the phenomena to be explained. A hypothesis must go beyond the phenomena so to speak, whilst it does not need to rely on available observation from other fields. The atomic hypothesis went beyond the facts of chemical combination then known, and assumed the existence of a certain type of system, atoms with special properties, to explain them. Atoms had not been observed, nor was it possible to identify them physically.

The hypothesis set out in what follows has many limitations, some of which are deliberate. This paper confines itself to performance of acts and learning in appetitive behaviour. Reasoning and avoidance behaviour are excluded as it is hoped to deal with these in a future paper.

Also, where undue complications of the hypothesis would have resulted from increased predictive power, simplicity has been preferred.

The paper begins by outlining the basic unit which may be supposed to underlie 'a simple act'. Then it describes the relations of the basic units among themselves. This is assumed in order to explain the performance of habits and instinctive behaviour. After this an explanation is put forward of how experience modifies this system to produce the phenomena of learning. Lastly, conflict is dealt with.

II. THE MECHANISM OF PERFORMANCE

There is no unit of behaviour but there is a unit mediating behaviour. Such a unit is presented by a list of the elements used in the system, together with statements concerning the relations of these elements among themselves. (Each numbered statement throughout the paper will describe a fresh feature of the system.)

There are five elements. These are called: (*a*) an analyser (a receptoral system); (*b*) a link; (*c*) a motor (or effector) organization; (*d*) an environment; (*e*) an internal medium.

There are three kinds of relation. These are called: (*a*) activating; (*b*) switching off; (*c*) causing to vary.

The elements are connected by the relations to form the basic unit in the following way:

(1) When the link is active, the motor organization is made to vary.
(2) The variations in the motor organization cause the environment to vary.
(3) A particular variation of the environment activates the analyser.
(4) The activated analyser switches off the link.

Activity in the link and so in the motor organization is terminated by the activation of the analyser attached to the link (see Fig. 1).

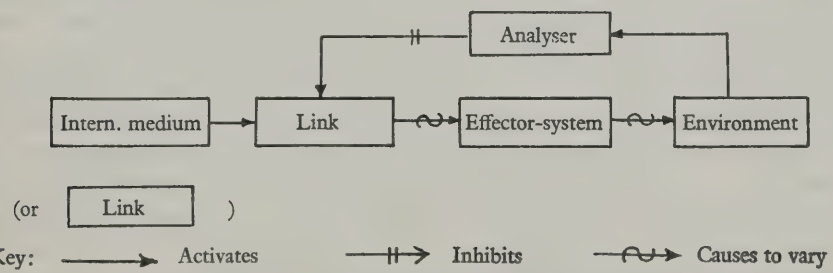


Fig. 1. *The unit.* The internal medium activates the link which causes the effector system to vary. This brings about changes in the environment which may eventually fire off the analyser attached to the link in this unit. The information from the analyser acts as feed back.

A 'stimulus' does not evoke a response, it terminates it. A dog will stop drinking when the receptors in its mouth and throat have been sufficiently stimulated (Bellows, 1939); a man will cease pronouncing a sound when he has heard himself doing so. Delay this information and he will stutter (Fairbanks & Jaeger, 1951). The question of the initiation of activity is therefore raised.

(5) One kind of link is set into activity by a specific or general feature of the internal medium.
(6) Another kind of link is activated by its neighbour.

No link is activated by both; either it is of the kind which is activated by the internal medium, or of the kind activated by another link.

Let us now expand the statements made and consider the units whose links are activated by the internal medium, applying the theory to some concrete examples. It has been stated that in some of these units, the link is set into activity by some specific change in the internal medium. This is the case in hunger, thirst, mating and instincts of a like sort. It is held that for each of these there exists a link, sensitive to a particular chemical state of the fluid surrounding it. The link responsible for male mating behaviour may be sensitive to male sex hormones: that concerned with thirst, cellular dehydration, and so on. The exact chemical identification of the states is, of course, unimportant to the theory *qua* system and is only tentative. Each of these links, it is held, is innately connected to a differing analyser or a set of analysers and so to differing sets of receptors. The link sensitive to cellular dehydration is in all probability connected to receptors in the mouth and throat, those mediating some of the special appetites to taste receptors, and the link mediating hunger mainly to receptors in the stomach. These identifications are again tentative. They appear plausible on what little evidence there is on this important subject.

(5*a*) The change in the internal medium irritates the link in proportion with the magnitude (or perhaps in some cases, rate) of this change.

The total amount of activity in the link will be a product of the amount of this change and the state of the link, which (4*a*) is determined by the amount and recency of the firing of the analyser innately connected to the link. The link sets the motor organization into activity. When the analysers connected to this system fire off, the irritability of the link to the particular chemical change which set it off begins to decrease. Therefore the time taken for the analyser to decrease the irritability of the link below threshold will vary with the amount of the change in the internal medium. The cessation of activity in the link will occur in spite of the continuance for some time or permanently, of the chemical state which initiated the activity, and this is the case in a normal animal. The mechanism of 'need' is therefore not a simple homeostat. However, in some of these 'needs', the particular chemical change which occasioned the activity which then led to a discharge in the appropriate analysers will be reversed as a remote consequence of the discharging of those analysers in an intact animal. For instance, in drinking, when the appropriate receptors have been stimulated to an extent sufficient to lead to a cessation of drinking, water must have entered the stomach. This water will then slowly redress the change which occasioned the drinking, comparatively long after drinking has ceased. If we assume that this type of system is in fact operating when we speak of the so-called instincts and physiological drives, we can explain the following observations.

The onset of these activities is not due (*pace* Hull, 1933) to intensity of stimulation. Selective abolition of afferent impulses has not been found to eliminate such activities (Tsang, 1938; Bash, 1939), nor has selective intensification of these produced an excess of activity (Montgomery, 1931). On the other hand, the onset of these activities appears to coincide with chemical changes in the blood (Beach, 1948) which Lashley (1938) concludes act centrally. It is then that such activities appear and persist until certain specific stimulation occurs. It has therefore been said that instinctive activity is caused by a stimulus deficit. (This looks paradoxical because the S-R theorists have long confused the two notions of stimulus and cause.)

When appropriate stimulation has occurred, activity ceases, even though the chemical change which occasioned the activity persists. A dog with an aesophageal fistula will drink the same amount as if water had entered its stomach (Bellows, 1939). The rats in Kohn's (1951) recent experiment stopped lever-pressing when a nutrient solution was inserted directly in their stomach, before it could have had an appreciable effect on the chemical changes which occasioned that activity. Further confirmation of Kohn's results and the interpretation put on them by the present theory is to be found in experiments performed by Miller & Kessen (1952), and also by Berkun, Kessen and Miller (1952).

This experiment also supports the present theory against that of Tinbergen (1951). Tinbergen believes that so much need is worked off or released in so much activity. After this release has taken place, activity ceases. But in this experiment no usual effector discharge occurred, and yet the animals appeared no longer to be hungry. It is plausible to believe, however, that certain receptors in the stomach had been stimulated, and this, according to the present hypothesis, is sufficient to explain why the animals were no longer disposed to eat.

The case of a dog with a full stomach which nevertheless drinks in gradually decreasing amounts after varying intervals (Bellows, 1939), might appear to contradict Kohn's (1951) results and with them the present theory. But the reader must be reminded that the dog was thirsty and the rats hungry, and that it has already been stated that the receptors remotely concerned with switching off thirsty behaviour lie in the throat and mouth, whereas those concerned with hungry behaviour lie mainly in the stomach.

The activity of the links, in some units (see Dennis, 1939) does not seem to rise or fall with fluctuations in some particular chemical state. Their activation seems to be due to some general feature of the constitution of the organism. Their existence is postulated because of the imperfect correlation, say, between the physiological needs and some appetites (Young, 1948) and because of certain preferences some of which might be called aesthetic.

Concretely interpreted, the hypothesis assumes that an animal, when we observe its performance, is searching for a given 'stimulus' or combination of 'stimuli' and that its performance ceases only when it has found them. When a man is searching for a shilling, he will ignore all 'stimuli' besides the shilling, and when he finds it his searching will cease. (This is of course an ideal case.) The activity ceases because certain specific analysers connected to the link causing the variations in output are fired off and switch off the link. The firing off of other analysers will be 'ignored'.

III. Behavioural sequences, innate and learned

To account for a whole behaviour sequence, such as the running of a maze, this hypothesis is extended in the following manner. The links in some units are exclusively activated by another link (see Deutsch, 1954).

(7) It is postulated that there is a single row of links, each link having only two neighbours, and the first and last only one, and that the first links (or the first few) are activated by the internal medium and these in turn excite the rest of the row of links which are only activated by an excited neighbour.

The first link is connected to the second, the second to the third, and so on. The first activates the second, the second the third. Therefore if the activity in the second ceases,

it also ceases in the third. The first link, however, will remain active, as follows from the above rule.

(8) Secondly, it is postulated that if all the links are active only the last one in the row (our number three) will vary output fully.

When activity in number three has ceased, number two will vary output. Number one's variation of the output will be almost completely inhibited when number three is varying output, but number two more partially so, as the inhibitory effects of each link on the output of the others are additive. Hence the system will display 'anticipatory' or 'preparatory' behaviour, a special type of which is the conditioned response. Anticipatory behaviour will appear only in certain conditions (as a corollary of a rule given later, which states that the link with a higher rate of firing will gain control of an output unit in case of competition for it). Hence the system would show such behaviour when the two links acted on different output units. For instance, a rat can run and munch and salivate at the same time, but it cannot run and lie down at the same time.

Thus we have two rules concerning the row of links. The first states that the row is activated from its first member through each member in turn. (The first member underlies the aim last in time to the observer, e.g. feeding in the goal box in the maze.) The second states that an active link partially inhibits variations in output in all the links through which it is activated and that this inhibition is additive (see Fig. 2).

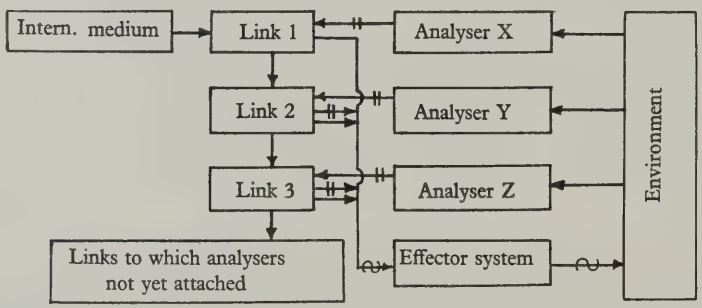


Fig. 2. *The organization of the chain.* The activation of a link activates the succeeding link. The activation of a link partially inhibits variation in effector system by preceding link. (Activation of a link only occurs when it has an analyser attached.) For key see Fig. 1.

We thus have a system which will pursue aims in turn, in a predetermined sequence. When it has run through the sequence, it will cease to be active. For when it has achieved its final aim (our number one in the sequence), there are no more activated links in the chain and therefore there is a cessation of variation of output. We have here a hypothesis which enables us to look in the same way at performance both instinctive and learned. Even if no new predictions could be made on it, it would still be of value as a co-ordinating and unifying hypothesis. Secondly, this system, owing to the first rule for the chain and the rule which states that a link is switched off when the analyser to which it is connected is touched off, will show a very interesting property. It will eliminate blind alleys or 'superfluous' moves, and break down on detour problems. This is deduced in the following way. Let us say that our row composed of three links is activated. Link number three will be varying output. If by chance the analyser connected to number two is fired off before the analyser connected to number three, link number two will be switched off, and with it link number three. Link number one will then vary output. The system will not

persist in pursuing an aim if that aim has been short-circuited, so to speak. If analyser number three is fired off by a cue up a blind alley, and analyser number two by a cue past the blind alley, the animal will not go up the blind alley, but pursue an aim nearer in order to the final goal. In some cases, that is in detour problems, the animal is allowed to see the goal but is not allowed to reach it except by pursuing aims which are now short-circuited by the firing off of analysers attached to links nearer in the chain to the first.

The activity in the row of links is transmitted by each member in turn from its first member. This first member is sensitive to changes in the internal medium and is set into activity by them, though its activity is cut short when the analyser which is connected to it is touched off.

(4*a*) Any link in the chain will take some time before it resumes full irritability after being switched off. It will be in a refractory phase.

Hence this first member will take some time before it resumes activity even though the internal medium may not change.

So far, it has been said, a system has been constructed consisting of five elements with three kinds of relations holding between them. The relations existing between them fully define the elements. The relations themselves can be given various physical interpretations, but they cannot be defined within the system. The psychological interpretation the system has been given has already been made clear both by the choice of words to label the various elements and relations and by the examples used to make the system clear to the reader. But one could equally well have chosen to interpret the system another way by giving it a meaning in terms of valves, relays or mechanical contrivances. In fact, the writer has built an electro-mechanical model embodying the main features of the present system, which is capable of learning mazes and discriminations, of eliminating blind alleys and of Maier-type reasoning (Deutsch, 1954).

IV. Learning

So far the connexions between the analyser, chain and effector system have been regarded as fixed. This may be the case in the system underlying instinctive behaviour. Here the variation in motor output may be largely absent too.

The system already outlined would not be capable of learning. The writer has supposed that there is a row or chain of links, each acting on others in the way described, with all the links already connected to analysers and output units.

(9) Attached to the end of each such a complete row are a great number of links also in a row, to which no analysers are yet attached.

(This means they will not be activated. A link to which no analyser is attached cannot be activated, according to a rule given later.)

(10) When the row is switched on from its first member many fresh analysers will be fired off in a novel situation before the analysers already attached to the row.

(11) When an analyser already attached to a link in the row is fired off the fresh analysers will be attached to the free links at the end of the row. The fresh analysers, which fired off nearest in time to the analyser already attached, will be fastened to the free links nearest in order to the links already occupied by analysers, so that temporal proximity of firing of analysers will lead to ordinal proximity of attachment on the row of links.

This is done by an ordering network acting in the following manner. Every analyser is connected to this network. As an analyser not already connected to a row of links fires, it

will be placed provisionally on the first free link attached to a row of links already occupied. If an analyser which is already occupying a link is fired off next in time to this fresh analyser, the fresh analyser will then be fastened to the free link on which it was placed. Next time this row is switched on, this previously free link will be the first to vary output. An animal will thus seek out the stimulation which occurred before the stimulation which it was seeking on a previous occasion (unless it is 'short-circuited', of course, in the same way as a blind alley).

If, however, the firing of an already attached analyser does not follow next in time after the firing of the analyser unattached to a link, but on the contrary the firing off of another unattached analyser occurs, no fastening to the first free link occurs in the way which has just been described. Instead the first free analyser to have fired off will now be displaced from the link on which it was placed (or provisionally attached) and shifted to the next link which is unoccupied. Meanwhile, the analyser which displaced it will provisionally occupy the first free link. If another fresh analyser fires off after this second one, the process of shifting goes on. The first fresh analyser to have fired will now be shifted to occupy the third free link provisionally, and the second analyser to the second link, which has just been occupied by the first analyser to have fired. The third and last analyser to have fired will now be connected to the first free link. This process of ordering and shifting will continue until the firing off of an analyser already permanently connected to the chain. This event will cause the permanent connexion to the row of those analysers which are provisionally connected to it. Their provisional connexion to their links now becomes permanent.

(12) When the row is again switched on, these previously free links will also be switched on and behave according to the rules already set forth.

Therefore the factor leading to reinforcement (to use the usual language) is the firing off of an analyser already attached to a row of links. (This is held to underlie the occurrence of the stimulation which the animal was seeking.)

An examination of experimental evidence by Sheffield will elucidate the hypothesis to the reader and will, it is hoped, provide a convincing explanation of it. Sheffield and co-workers (Sheffield & Roby, 1950; Sheffield, Wulff & Backer, 1951) have found that copulation without ejaculation is a primary reinforcing factor in rats or, to put it in another way, possesses reward value which does not diminish with time. They have also found that the taste of saccharine, though non-nutritive, appears to have primary reinforcing properties. These findings constitute a puzzle for those who say that tension reduction or biological need reduction leads to reinforcement. On the present hypothesis the findings form no exception and could have been predicted. As the rat has an unlearned preference for saccharine and aims at copulation without previous accidental experience of it, it is inferred that the analysers signalling a sweet taste (or the various stages of copulation) are connected permanently and innately to the first part of a row of links. This being the case, any analysers fired off previous to these will be connected to the free part of that row of links. Next time the row of links is activated the previously free part will be activated too, with the consequence that an animal will now seek out not only stimulation (or stimulus-configuration) innately attractive to it, but also that stimulation which occurred previous to this stimulation on a past occasion. Thus learning will have taken place in the same manner as it is always supposed to occur on the present hypothesis.

This system converts a temporal series of events into a receptoral series. Whenever the system secures any aim, it acquires additional aims in a novel situation. For those analysers, which were touched off previous to the firing off of connected analysers, and were not themselves already connected there, are retained, connected to the chain.

This may be expressed in another way. Learning, in one of its aspects, is an ordering of experience. This feature is expressed in the system by connecting the analysers to an ordering network, which joins them to the free part of a row of links in reverse order as they were touched off. The analysers ordered on the chain will be connected and fixed to it whenever the firing off of an analyser attached to an active row occurs. That is, they will not now be ordered and pushed down any more, but remain stably connected whenever an analyser already connected is touched off and thus switches off its link. The firing off of the analyser of the first unit in the chain would bring this about in the first instance. The system would continually be installing 'new aims'. The attaining of these in a fixed environment will enable the system to reach its final goal (represented by the firing off of the analyser of the first unit) with certainty and will lead to a pattern of activity which is flexibly fixed (by the cues the system is looking for). But this activity will not merely be a repetition of the first random trial. For the system will eliminate blind alleys and false moves, a property of this system which has already been deduced.

It follows that the amount of learning in this system is a function of the amount already learnt, given an environment with a stable composition. A young animal will learn more slowly than an older one. Similarly, an animal with a less varied experience will also be slower. This is pointed out by Hebb (1949). It is predicted by the present theory that the same will hold true for latent learning. Some kind of vicarious trial and error would also be predicted. The animal on this hypothesis is looking for its stimuli.

This view obviates the necessity of distinguishing between primary and secondary reinforcement. The problem of removing the distinction therefore does not arise.

V. Quantitative consideration

So far many of the relations among the elements of the system have been all-or-none in nature with no quantitative or temporal variation. The system set up is, however, complete as it is. It could be given a certain kind of physical interpretation. But to serve as a calculus from which to predict animal behaviour these relations will be postulated to vary with time and in quantity.

Let us formulate the relations in the system in this way:

(A) The higher the activation of the internal medium, the higher the activation of the first link (5*a*).
(B) The higher the activation of the first link, the higher the activation of the other links (6*a*).
(C) The higher the activation of the links, the higher the activation of the effector system (1*a*).

The system would thus act with greater intensity and speed with increased 'need'. If we wish to give a neurological interpretation to the system, intensity can be thought of as frequency of firing along the connexions which mediate the relations between the elements.

(D) All the consequences of the touching off of an analyser already connected to the chain will increase with the following:
  (*a*) the amount of firing from the analyser;
  (*b*) the degree of activation of the link switched off.

Therefore, (i) the length of the refractory period (a state which gradually returns to normal) of the link will vary with these factors (4*a*), and (ii) the threshold of a link's activation

by its predecessor will decrease with the amount of firing from the analyser of the predecessor and the degree of activation of the predecessor and the number of times the predecessor has been switched off by its attached analyser, following the link.

(E) A link will be unexcitable by its predecessor unless it has an analyser attached. Its threshold will then vary as stated mainly in D(i) and F(13). (The concept of threshold is here used in a quasi-physiological sense.) The actual amount of activation in a link will be a function of the amount of activation of the link activating it and its own threshold.

(F) A rise in this threshold of the link for activation will occur when the next (preceding) link is varying output. The rise will increase with the length of the period of output by the variation of the next link (13).

This will occur when the feed back required to switch off the next link does not arrive immediately. The threshold will rise with time until the link cannot be activated by the next (preceding) link (14).

(G) Both these effects of lowering and raising the threshold will fade out with time, leading to the phenomena of forgetting and spontaneous recovery (15).

This statement of the relations in the system as continuously variable gives the theory a power of quantitative prediction. Unfortunately, within the limits of a single paper, some only of the empirical consequences can be put forward.

Massed learning will be less effective than spaced learning, but massed extinction more effective than spaced extinction (see especially section D).

Furthermore, if we agree that the link which gains control of output is always the one with a lower threshold (this is explained later), this will lead to the selection in a choice situation of the most often rewarded habit and one which has been acquired under the greatest need (see D).

Also, these rules will lead the system to select the most constant cues in the environment. Kinaesthetic cues are therefore ordinarily of minor importance except when they are generalized, but distant cues are more important because of their relatively constant nature.

One of the consequences of the hypothesis which it is now possible to draw is that extinction and reinforcement of a habit (or response) can occur without its performance. The application of the theory in this respect will be made clearer if we use it to interpret a study of extinction by Seward & Levy (1949). In this two groups of rats, E and C, ran down an elevated pathway from a starting platform to a goal platform for a food reward. After some training each rat in group E was placed on the goal platform with no food on it. The C group was identically treated except that it was placed on a strange platform. 'Five such treatments were given. . . .On the following day all rats were given extinction trials to a criterion of two successive refusals to leave platform A (the starting platform) within three minutes.' 'Group E reached the criterion in a mean of 3·12 trials, group C in 8·25; the difference was significant at the 0·01 level.'

During the learning trials, it is assumed, a row of links becomes attached to the analysers which fired off before the rat found food. Nearest on the row to the analysers fired off in the presence of food would be analysers firing off in the presence of the goal platform. When the rat finds itself on the empty goal platform in a hungry state the appropriate link (or links) are switched off and the next (preceding) 'food' link begins to vary output without being switched off. This increases the activation threshold of the 'goal platform' link according to paragraph F in this section. Therefore the links which are activated through

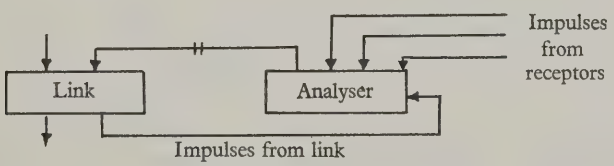
his link will also be activated to a lesser degree. Hence the 'habit' will run off at a ecreased speed (section B and C) and show extinction with repetition of this procedure. habit can thus be weakened without actually having been elicited.

VI. ANALYSERS AND OUTPUT

he hypothesis put forward does not claim to give an explanation of the way the total vailable information is sorted. It takes such things as the perception of configurational roperties as given. What it does try to explain is how information influences the actual nd future performance of the animal. If one were to construct a machine model of the ystem, any simple device which delivered a pulse when set off could be equated with an nalyser as defined in the system. Equally well, a most complex device could be used. s long as we have a device which is set off only by a specific feature or group of features n the environment it can be used as an analyser in the system. The system as put forward an, however, account for certain characteristics of perception. The analyser and link to hich it is connected interact, in such a way that a firing off of the analyser switches off link (a rule already made).

(16) Some of the excitation in a link is transmitted to the analyser, increasing with the number of times the analyser has switched off the link, thus lowering the analyser's threshold for firing accordingly.

he analysers have the observed property of firing off more readily the more frequently ney have already been fired. Secondly, they also fire off more readily if the chain to which ney are attached is activated. That is, the animal or human will tend to see more readily nat for which it is looking. By 'more readily' is meant 'on less complete information nan that which is necessary to fire the analyser off when these conditions of experience nd motivation are absent' (Fig. 3). This way can be deduced when we consider that the



g. 3. *The relations between analyser and link.* The impulses from link and receptors additively cause firing off of an analyser. The activation of the link is a function of the frequency of past firings off of the analyser attached to it, and the degree of activation by the internal medium or preceding link. Therefore motivation and experience will influence perception. For key see Fig. 1.

citation in a link is a function of the excitation of the previous link and the threshold of e link itself. The threshold of the link itself is a function of the number of times the ing off of the attached analyser has occurred. If we treat the impulses from the link to e analyser and the incoming impulses to the analyser as additively causing the firing of e analyser, then it follows that the more frequently the analyser has been fired off owering the threshold of the link) and/or the greater the excitation of the link by previous nks, the fewer incoming impulses are necessary to fire it off (see Fig. 3). Thus the influence experience and motivation is explained by the hypothesis. It can be deduced that an imal in a high state of need will eat things which it previously would not have regarded food, and that it will indulge in 'vacuum activity' (to use Tinbergen's phrase (1951)). equently experienced figures will tend to be 'completed'. Hungry men will tend to see

the environment in terms of food. Ambiguous information will be 'interpreted' in accor dance with need and experience.

(17) The output which occurred just before the firing off of an analyser fixed to a link will be ther when the link is activated again.

If in searching for a certain station on the wireless one switches off as soon as one ha found it, the next time one switches on, one will not have to search again, unless it ha shifted.

VII. Conflict and repression

Though the hypothesis advanced in the previous pages is capable of explaining th learning and performance of a single habit, it must be expanded so that it may be applie to situations in which there is a choice. Let us therefore assume that there are man chains, with the properties described in a system, all of which may control output. N further rule would need to be incorporated if all these were assumed to activate outpu in turn. But sometimes the system may be confronted with a choice.

(18) The rule which is therefore added is that if there are two links competing for the same output, th one with a higher rate of firing will secure output.

Now the rate of firing in an activated link has already been made to depend on th degree of activation of the link activating it and the threshold of the link being activated One of the factors affecting the threshold is the temporary rise in threshold (the refractor period) in a link after the firing off of the attached analyser. Therefore, if an animal i trained equally on two habits of equal length and equal reward, and then allowed t choose between them on closely spaced successive trials, then it will alternate regularl between them. This deduction is confirmed by the experimental work of Dennis (193 and Heathers (1940).

Hull would, of course, predict the same (from the postulates concerning reactiv inhibition). On the other hand, he would predict turn alternation, whereas the presen theory in the usual situation demands what might be called stimulus, place or goa alternation. The refractory period of a link is due to the firing off of the attached analyse This leads to a temporary weakening of the 'attractiveness' of those parts of the enviro ment which have fired this analyser. An experiment performed by Montgomery (1952) t test these alternative possibilities supports unequivocally the deduction from the prese hypothesis.

The slower firing link will still give rise to output by varying effectors which the faste firing link is not varying. It will go on varying output in this way until its attache analyser is fired off. Not being able to use all the possible effectors, it will keep active f a longer time than if it could control them entirely. This consequence of the hypothes may be used to explain psychosomatic symptoms. Let us take repressed anger. It assumed that a part of the output of the innate anger unit is fixed (at least as far as t autonomic part is concerned). If angry behaviour is unrepressed, the firing off of a attached analyser switches off the unit and the somatic reaction is transient. This is would be predicted. But where angry behaviour does not occur when the anger unit activated, because of faster-firing units which control most of the effectors, it will st bring about an autonomic anger reaction. But this is not transient. It often assumes a almost chronic character. Why the repressed 'affect' should differ in this way from t

expressed is left a puzzle on most theories. On the present hypothesis the answer is quite simple. The unit persists in activity until the component analyser is fired off. In the case of repression of output, it will not be. Hence the persistence, hence the ulcers.

Other kinds of behaviour resulting from conflict would also be predicted. Restlessness, oscillation between alternatives, would be expected under some conditions. In the complete system set up, many rows of links will be in varying degrees of activity. Those in which activity is strongest will determine behaviour.

For example, it has been demonstrated by Hull (1951) and Leeper (1935) that rats can be taught to take one path to water when thirsty and another to food when hungry. Hull explained this by postulating two diverse persistent stimuli, characteristic of each drive, separately conditioned to which the animal would learn to make a different response. Kendler (1946), however, trained rats in a simple T-maze when they were simultaneously hungry and thirsty. There was water in one goal box and food in the other. All animals had equal opportunity to explore both goal boxes. In the test series the rats were made either hungry or thirsty, not both. They were able to make the appropriate choice. This makes a difficulty for Hull. There was only a single pattern of drive stimuli during training, generated by the state of being simultaneously hungry and thirsty. The supposition that rats learn to respond differentially to diverse patterns of drive stimuli in making the choice is made unlikely.

The present hypothesis, on the other hand, presents another way of looking at this evidence. There are two rows of links, one with analysers (presumably in the mouth and stomach, as suggested above) which fire off when food is ingested, and the other when the water is ingested. These analysers are attached to the first link, which is excited by certain changes occurring around it. The one to which the 'food' analyser is innately attached is sensitive to changes occurring when there is a nutritional deficit, the other to which the 'drink' analyser is attached, is excited when there is a water shortage. Now when either or both these links are excited, and the animal is thrown into activity, many analysers unattached to either of these two rows of links will be fired off. If, say, the already attached 'drink' is fired off immediately after these, they will be connected to the links following this attached 'drink' analyser, and not the row of links following the other innately attached 'food' analyser. Next time the link to which the innately connected 'drink' analyser is connected is excited by the appropriate chemical change, the row of links following it will also be excited. These will now have attached to them the analysers which were fired off last time before the innately attached 'drink' analyser, whereas the row of links to which the 'food' analyser is fixed has not. (This hunger-excited row of links may also have some analysers attached to it if some fired off prior to the finding of food.) Therefore when the animal is thirsty, it will begin to look sequentially for those stimuli which occurred prior to its finding water. The same of course applies, *mutatis mutandis*, when it is hungry. Therefore the animal will select that sequence which when it was 'hungry' led to a 'food' result when it is hungry again, on account of the mechanism outlined. It clearly does not make a difference to this ability whether the animal was thirsty or not when it was hungry and found food.

The firing off of an attached analyser can occur to a link even though it is not determining activity, as follows from the basic rules of the system. The firing off of an attached analyser varies its consequences according to the level of activation of the links in which it occurs. Therefore latent learning would be less efficient. The firing off of an attached

analyser according to the hypothesis, is responsible for 'reinforcement'. Without th firing off of an attached analyser no learning will take place. Thus it is predicted that i an animal is placed in a maze hungry, and its physiological deficiency is artificially corrected in the goal box, this need reduction will not act as 'reward'. For what leads t learning is the touching off of analysers which switch off activity in a link. Thus the tast of saccharine acts as a reward, whereas glucose, untasted and unfelt in the stomach, it i predicted, will not.

VIII. Summary

A hypothesis to explain the phenomena of appetitive behaviour has been put forward This hypothesis is a formal system which is given such an interpretation that proposition concerning appetitive animal behaviour follow from it. It can equally well be given othe types of interpretation. For instance, an electro-mechanical learning and reasonin machine has been made by the writer embodying the system (Deutsch, 1954). A stimulu shuts off activity and does not initiate it. The initiation of activity is attributed to centre which are affected by physical states within the body, without the mediation of afferen pathways. Evidence concerning physical needs and instinctive activities is shown to fi such a view.

To account for sequential activities a network is suggested, in which a particula physical state selects a series of receptors which will switch off various types of activit in turn. An explanation of the elimination of blind alleys and superfluous moves is give in terms of the characteristics of this network.

Learning is viewed as a conversion of temporarily successive events into receptora series, arranged in that order in which stimulation occurred. This series may be maintaine at the end of another receptoral series, under certain simple conditions. 'Reinforcement' according to the hypothesis, is due to the occurrence of that stimulation which the syste was selecting. The distinction between primary and secondary reinforcement is thu rendered superfluous. This view is applied to recent evidence on reward. Deduction concerning the role of experience and motivation in perception are also made.

Conflict and choice situations are viewed as the simultaneous operation of many suc systems. Since the selection of a particular sequence of activity occurs by and throug the units underlying the goals of the organism, some experimental evidence on th selection of habits by their 'anticipated result' is explained.

References

Bash, K. W. (1939). An investigation into a possible organic basis for the hunger drive. *J. Comp. Psycho* xxviii, 109–34.

Beach, F. A. (1948). *Hormones and Behaviour; a Survey of Interrelationships between Endocrine Secretio and Patterns of Overt Response.* New York: Paul B. Hoeber.

Bellows, R. T. (1939). Time factors in water drinking in dogs. *Amer. J. Physiol.* cxxvi, 13–19.

Berkun, M. M., Kessen, Marion L. & Miller, N. E. (1952). Hunger-reducing effects of food by stomac fistula versus food by mouth measured by a consummatory response. *J. Comp. Physiol. Psycho* xlv, 550–5.

Dennis, Wayne (1939). Spontaneous alternation in rats as an indication of the persistence of stimul effects. *J. Comp. Psychol.* xxviii, 305–12.

Deutsch, J. A. (1954). A machine with insight. *Quart. J. Exp. Psychol.* (in the Press). To be publishe in Jan. 1954.

Fairbanks, G. & Jaeger, R. (1951). A device for continuously variable time delay of headset monitorin during magnetic recording of speech. *J. Speech Dis.* xvi, 162–6.

HEATHERS, G. L. (1940). The avoidance of repetition of a maze reaction in the rat as a function of the time interval between trials. *J. Psychol.* x, 359–80.

HEBB, D. O. (1949). *The Organization of Behaviour; a Neurophysiological Theory.* New York: Wiley.

HULL, C. L. (1933). Differential habituation to internal stimuli in the albino rat. *J. Comp. Psychol.* XVI, 255–73.

HULL, C. L. (1951). *Essentials of Behaviour.* New Haven: Yale University Press.

KENDLER, H. H. (1946). The influence of simultaneous hunger and thirst drives upon the learning of two opposed spatial responses of the white rat. *J. Exp. Psychol.* XXXVI, 212–20.

KOHN, M. (1951). Satiation of hunger from food injected directly into the stomach versus food ingested by mouth. *J. Comp. Physiol. Psychol.* XLIV, 412–22.

LASHLEY, K. (1938). The experimental analysis of instinctive behaviour. *Psychol. Rev.* XLV, 445–71.

LEEPER, R. (1935). The role of motivation in learning: a study of the phenomenon of different motivational control of the utilization of habits. *J. Genet. Psychol.* XLVI, 3–40.

MILLER, N. E. & KESSEN, MARION L. (1952). Reward effects of food via stomach fistula compared with those of food via mouth. *J. Comp. Physiol. Psychol.* XLV, 555–64.

MONTGOMERY, K. C. (1952). A test of two explanations of spontaneous alternation. *J. Comp. Physiol. Psychol.* XLV, 287–93.

MONTGOMERY, M. F. (1931). The role of salivary glands in the thirst mechanism. *Amer. J. Physiol.* XCVI, 221–7.

SEWARD, J. P. & LEVY, N. (1949). Sign learning as a factor in extinction. *J. Exp. Psychol.* XXXIX, 660–8.

SHEFFIELD, F. D. & ROBY, T. B. (1950). Reward value of non-nutritive sweet taste. *J. Comp. Physiol. Psychol.* XLIII, 471–81.

SHEFFIELD, F. D., WULFF, J. J. & BACKER, R. (1951). Reward value of copulation without sex drive reduction. *J. Comp. Physiol. Psychol.* XLIV, 3–8.

TINBERGEN, N. (1951). *The Study of Instinct,* Oxford: Clarendon Press.

TSANG, Y. C. (1938). Hunger motivation in gastrectomized rats. *J. Comp. Psychol.* XXVI, 1–17.

YOUNG, P. T. (1948). Appetite palatability and feeding habit: a critical review. *Psychol. Bull.* XLV, 289–320.

(*Manuscript received* 12 *January* 1953)

ELECTRICAL RESPONSES OF THE BRAIN IN RELATION TO BEHAVIOUR

By A. C. MUNDY-CASTLE

National Institute for Personnel Research, South African Council for Scientific and Industrial Research, Johannesburg

I. *Introduction* (p. 318). II. *Flicker response and behaviour deviation* (pp. 318–321). III. *Differential aspects of the flicker response* (pp. 321–325): (i) *Epileptiform* (pp. 321–322); (ii) *Psychopathological* (p. 322); (iii) *Emotional* (p. 322); (iv) *Sensory* (p. 323); (v) *Intellectual* (pp. 324–325). IV. *Flicker response and cortical excitability level* (pp. 325–327). V. *Summary and conclusions* (pp. 327–328). *References* (pp. 328–329).

I. Introduction

Whilst there is a large body of knowledge concerning the clinical significance of abnormal electroencephalograms (EEG's), relatively little is known of the role of the electrical rhythms of the brain in relation to normal behaviour. A recent development in clinical electroencephalography has been the move from static to 'functional' application, in which analysis is made of EEG changes during varying conditions of cerebral function. These variations can be produced by a number of methods, e.g. drugs, artificial sleep, insulin hypoglycemia, hydration, in fact anything likely to reveal differences in the brain's capacity to withstand controlled stress. These methods were first utilized for clinical purposes to establish EEG thresholds in patients suspected of epilepsy, and the results have been highly successful. Subsequent workers have utilized such methods for controlled scientific research (Gastaut & Hunter, 1950; Hill, Loe, Theobald & Waddell, 1951). A form of environmental control not strictly biochemical and consequently more simple to apply is rhythmic photic stimulation. Adrian & Matthews (1934) first reported that observable responses could be evoked in the EEG by flashes of light in the eyes, whilst in 1946, Walter, Dovey & Shipton (1946) confirmed that seizures could be elicited by such rhythmic stimulation. This development arose out of Walter's (1946) suggestion that seizure wave-forms might be built up by the synthesis and synchronization of a number of harmonically related rhythms. It was argued that this could be done by augmenting specific cerebral rhythms by photic stimulation. This was shown to be so, and the method is now used as an activation procedure in numerous EEG departments.

The present paper deals with the psychological implications of EEG changes during rhythmic photic stimulation. The clinical aspects have been discussed at length in numerous other publications (Gastaut, 1949, 1950; Mundy-Castle, 1952*b*; Walter & Walter, 1949*a*, *b*; Walter, 1950*b*).

II. Flicker response and behaviour deviation

Since the clinical aspects of responses to photic stimulation are not relevant to the present discussion, it need only be mentioned that approximately 25 % of epileptics show abnormal responses (Gastaut, 1949; Mundy-Castle, 1952*b*; Walter, 1950*b*). But a finding from our laboratory which seems of considerable psychological importance occurred during a recent study of the clinical significance of photic stimulation (Mundy-Castle, 1952*b*).

A Scophony electronic stroboscope was used in conjunction with a six-channel Ediswan electroencephalograph. This stroboscope is used for all our flicker studies. It gives a blue-white flash with a time-constant of 15 μsec. and intensity of approximately 88,000 candles; the flash frequency can be varied from 3·5 to 100 f./s., with 3·5–30 f./s. being the most effective. Stimulation is usually through closed lids with the lamp fixed about 9 cm. from the subject's eyes. The effects of eye opening and closure are also always studied. Flashes are recorded on the EEG by means of a photocell resting on the subject's forehead. For full details of recording see Mundy-Castle (1953). Among a large group (290 subjects) of probable epileptics who were tested by this method, there was a group of nineteen behaviour deviate children with fits and a group of seventy-seven probably epileptic children without behaviour deviation. The incidence of abnormal responses to photic stimulation in these two groups is compared in Table 1, together with the incidence of abnormal responses in another group of twenty-two behaviour deviate children without fits.

Table 1. *Flicker response abnormality*

(Children, clinical groups)

	Incidence of abnormality		Age	
	N	%	*M*	S.D.
(1) Behaviour deviates with fits ($N=19$)	13	68	10·79	3·05
(2) Behaviour deviates without fits ($N=22$)	8	36	10·09	5·25
(3) Children with fits only ($N=77$)	19	25	9·87	4·35

Significance of differences in incidence of abnormal response:

(1) + (2), $t=2·17$, $P<0·05$, significant
(1) + (3), $t=3·73$, $P<0·01$, significant
(2) + (3), $t=1·03$, $P>0·05$, not significant

It will be seen that behaviour deviation, whether accompanied by fits or not, gives a greater incidence of abnormal flicker response than fits alone. All these incidences are significantly different from that in a normal adult group ($N=154$, % $A=3·90$, Mundy-Castle, 1953). Unfortunately, a normal child group has not been available, but out of twenty children with physical developmental errors, 15% showed abnormalities, and there is little doubt that the incidence would be lower still in a non-clinical group.

These figures suggest that an abnormal flicker response may in some way be related to behaviour deviation. To elucidate this possible relationship we must consider the construction of an EEG flicker response pattern. It appears that electrical rhythms in the brain can be initiated or augmented by a process similar to resonance; in other words, if an area of the brain is subjected to rhythmic impulses corresponding to its own latent or actual frequency, it may itself oscillate for as long as stimulation is maintained. A further apparent property is harmonic resonance: the area may be activated by stimulation at a frequency harmonically related to its own: for example, 12 c./s. responses may occur at 4 f./s., this being the 3rd harmonic of the flash rate, or alternatively 3 c./s. responses may occur at 9 f./s., this being the 3rd subharmonic of the stimulus frequency. Such subharmonics are rare in normals. The variety of possible response patterns which can be evoked by rhythmic photic stimulation becomes very considerable when it is realized that all frequencies between 3 and 30 c./s. may be effective, and each may evoke a number of harmonic or subharmonic components. Furthermore, the response pattern may vary according to location, so that, for example, the occipital regions may show 2nd

and 3rd harmonics, the temporal areas the 1st harmonic (fundamental), and the frontal areas the 3rd and 4th subharmonics. In addition, the distribution and nature of evoked activity will vary according to the stimulus frequency and the state of the subject. This will be discussed later.

There are three main types of abnormal flicker response: frontal spike, wave-spike and dysrhythmic (Mundy-Castle, 1952*b*). One of the most significant features of the last two is the predominance of high voltage slow subharmonics of the stimulus frequency, these being attributed to subharmonic resonant excitation. When such slow waves exceed a certain amplitude and adopt a specific phase relationship with other higher frequency components of the response pattern, the conditions are set for the appearance of typical epileptic patterns, specifically the wave-spike formation. Such a response is rarely confined to the occipital areas and is usually accompanied by observable behavioural changes, either sensory, motor or both. Gastaut & Hunter (1950) have shown that the abnormal response occurs as a result of a 'short-circuiting' of afferent impulses at a thalamic level, these irradiating to excite other than visual areas of the brain. The leakage is attributed to abnormally low thresholds within the thalamic system. It has been postulated that the subharmonic responses evoked by flicker may bring about a lowering of these thresholds, and that such subharmonic resonant excitation reflects a state of instability within thalamo-cortical circuits (Mundy-Castle, 1952*b*; Walter & Walter, 1949*b*).

The subject may be an epileptic, in which case some kind of epileptic fit occurs. Or he may be normal, psychoneurotic, psychotic or a behaviour deviate, the behaviour associated with the abnormal flicker response varying widely within all these categories. It was found (Mundy-Castle, 1953) that out of 154 normal adults, the six who displayed an abnormal response showed definite personality deficiencies or signs of social or occupational or emotional instability. A similar trend was observed with those subjects showing questionable responses, although to a lesser degree. The abnormal flicker response is therefore suggestive of a constitutional disposition to some form of instability not necessarily epileptic, although possibly closely related. It may be that this instability is a form of cortical immaturity.

This postulate of cortical immaturity is comparable to that of Hill (1944) concerning the excessive theta rhythm (4–7 c./s.) found in some aggressive psychopaths; a finding which is attributed to a delayed or impeded maturation of cortical control over subcortical structures. His hypothesis receives some support from findings in the three clinical groups of children now reported on; among these groups the mean amplitude of responses within the theta band showed a distribution similar to that of abnormal flicker responses. The figures are shown in Table 2.

It is also significant that extreme aggressiveness was reported in 63% of the combined behaviour deviate groups whilst only 9% of the children with fits only were reported to have quick tempers or aggressive episodes.

We have, therefore, tentatively concluded that central responses to photic stimulation in children may reflect to some extent the ability of the child to withstand environmental stresses of both a physical and psychological nature: thus incomplete social adjustment, if due to a constitutional instability or immaturity within the central nervous system, may be paralleled by a breakdown at the electrochemical level during rhythmic physical stimulation, the result being an abnormal EEG flicker response.

Whilst it is probable that the abnormal flicker response is constitutionally determined (Daly & Bickford, 1951; Fairweather, O'Sullivan & Walter, 1949), there is some evidence to suggest variations in its manifestation in the same individual. One of the 6/154 normal young adults showed an abnormal flicker response at the beginning of stimulation only, it being impossible to re-elicit this thereafter. Similar response-sequences have occurred in both senile and manic-depressive psychotic patients (unpublished observations), whilst Walter (1950*a*) has also reported such cases. But occasional psychoneurotic patients

Table 2. *Theta response amplitudes*

(Children, clinical groups)

	Theta response amplitude (μV.)	
	M	S.D.
(1) Behaviour deviates with fits	40·86	15·70
(2) Behaviour deviates without fits	33·50	17·77
(3) Children with fits only	28·05	17·01

Significance of differences between mean theta response amplitudes:
(1) + (2), $t = 1{\cdot}41$, $P > 0{\cdot}05$, not significant
(1) + (3), $t = 3{\cdot}13$, $P < 0{\cdot}01$, significant
(2) + (3), $t = 1{\cdot}28$, $P > 0{\cdot}05$, not significant

have displayed marked abnormalities on several recordings, whilst their disturbed psychological state has persisted, but normal responses when it has been alleviated (Mundy-Castle, 1951*a*, 1952*b*). It is concluded that variations in intrathalamic thresholds sometimes occur during psychoneurotic states, and that these can be raised to a 'normal' level by successful treatment. Further study is needed to reveal both the way in which such treatment operates and the relationships between the lowered thresholds and the neurosis. It should be noted that in some epileptics the abnormal flicker response is removed by suitable drug therapy, evoked wave-spike responses of the *petit mal* class being particularly susceptible to Tridione. Gastaut & Hunter (1950) have shown experimentally that this is specifically due to a raising of intrathalamic thresholds prohibiting the abnormal leakage of afferent impulses.

III. Differential aspects of the flicker response

There are several aspects of the flicker response apart from this apparent relationship between abnormality and disposition to behaviour disturbance. We have found it helpful to classify these aspects as follows: (1) epileptiform; (2) psychopathological; (3) emotional; (4) sensory; (5) intellectual.

All of these may co-exist in the same individual. They will be considered in detail.

(i) *Epileptiform*

This has already been mentioned, and deals simply with abnormal responses of epileptic appearance. We have purposely used the term 'epileptiform' since not only epileptics show an abnormal response and there is no justification for assuming that those normals, psychoneurotics, psychotics, behaviour deviates, etc., who display such a response are 'latent' epileptics. Nevertheless, the nature of the wave-forms, their occurrence and their relationships to one another are similar to or indistinguishable from

epileptic activity and overt seizures may also be evoked. Since the problems of the abnormal response are largely of neuropsychiatic interest there is no need to discuss them further here. Detailed discussions are available in papers by Gastaut (1949, 1950), Mundy-Castle (1952*b*), Walter & Walter (1949*a*, *b*) and Walter (1950*b*).

(ii) *Psychopathological*

The necessity for the remaining four classes arises since flicker responses may be accompanied by widely differing subjective and behavioural changes according to response frequency, complexity, location and the nature and state of the subject at the time of the test. Broadly speaking, psychopathological refers to delusional or hallucinatory states directly evoked by flicker, or to pathological fear, anxiety or other similar reactions which can be shown to be related to evoked electrical responses. Such reactions may persist for some time after cessation of stimulation, but are fortunately quite rare. Conscious attitudes, memories and dream-states have also occasionally been reported. The very nature of these responses, combined with their infrequency, renders a controlled study impossible, but less dramatic and less painful responses have been studied fairly carefully as will be discussed below. Gross EEG flicker response abnormalities, accompanied by extraordinary subjective sensations, are attributed to activation of circuits involved in the storage of memories and in emotional expression, these involving both subcortical and cortical structures, the break-through which precipitates their release occurring primarily at a subcortical level.

(iii) *Emotional*

The psychopathological category grades over into the emotional, the latter referring to affective changes not unduly excessive and not persisting after cessation of stimulation. Such changes are usually accompanied by delta (1·5–3·5 c./s.) and theta rhythms from within the fronto-temporo-parietal regions, these either directly evoked by flashes at corresponding frequencies or as subharmonic or near subharmonic components of the stimulus frequency. Emotional responses not falling in the psychopathological category are fairly common in normal subjects and usually consist in mild anxiety, anger or confusion (Fig. 1).[1] Such responses may behave in both a regenerative and degenerative fashion: stimulation of a subject in a bad mood is more likely to result in augmented theta rhythm and anger than if he is in a good mood, whilst introspective analysis of one's feelings, or prolonged stimulation at the same frequency, may diminish both the evoked rhythms and the strength of the affective reaction (Fig. 2). Subjects have also reported that with practice they are able to dissociate themselves from the light flashes and the resulting disturbing sensations by deliberately thinking about other things. A change in stimulus frequency, creating unfavourable conditions for the evocation of theta or delta rhythm, will also remove the emotional components of the response. The conclusions from these experiments support earlier findings concerning spontaneous theta rhythm (Mundy-Castle, 1951*b*; Walter, 1950*a*) namely that it tends to accompany unpleasurable emotional expression. It is most probably mediated by thalamo-cortical pathways with a subcortical origin.

[1] These figures show superimposed automatic frequency analyses. The heights of the vertical deflexions reflect the relative amplitudes or amounts of specific frequencies over repeated 10 sec. epochs. The relevant frequencies have been indicated. The channel being analysed is shown by the number of deflexions in the topmost line.

(iv) *Sensory*

Rhythmic photic stimulation may also evoke sensory patterns, these being attributed both to interaction between spontaneous and evoked electrical rhythms and to the augmentation of activity harmonically related to the stimulus frequency. The most common are visual, usually in the form of moving concentric rings with a faint cross or regular spoked figure radiating from the centre, together with myriads of small shadows or criss-crossed elements forming a shifting network of colours changing both spontaneously and with different stimulus frequencies. At low frequencies (below 8 f./s.) there is less movement and vivid colour seems to fill the head, whilst above about 30 f./s. a homogeneous blue-white field is seen. Every subject we have tested (approximately

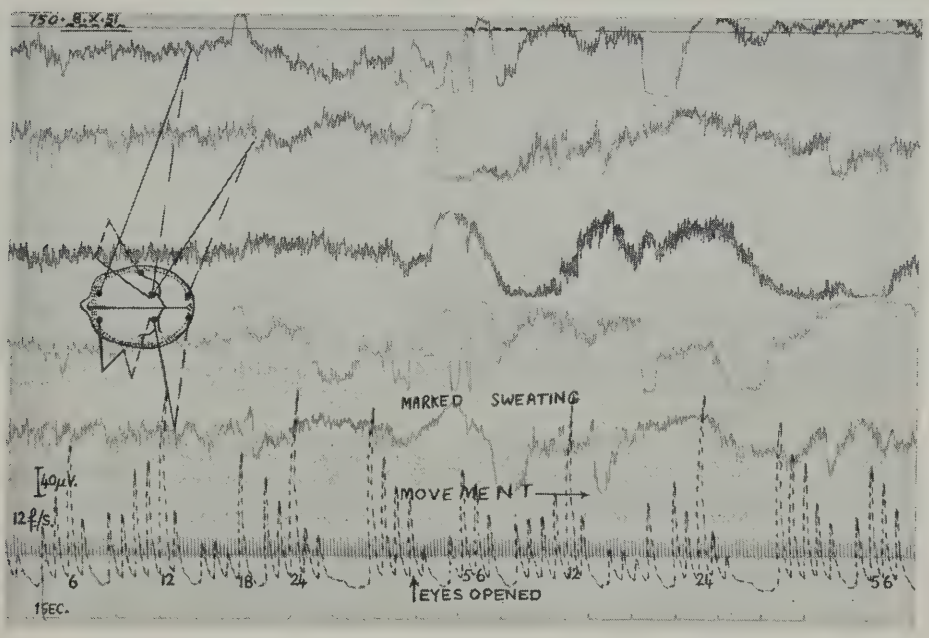
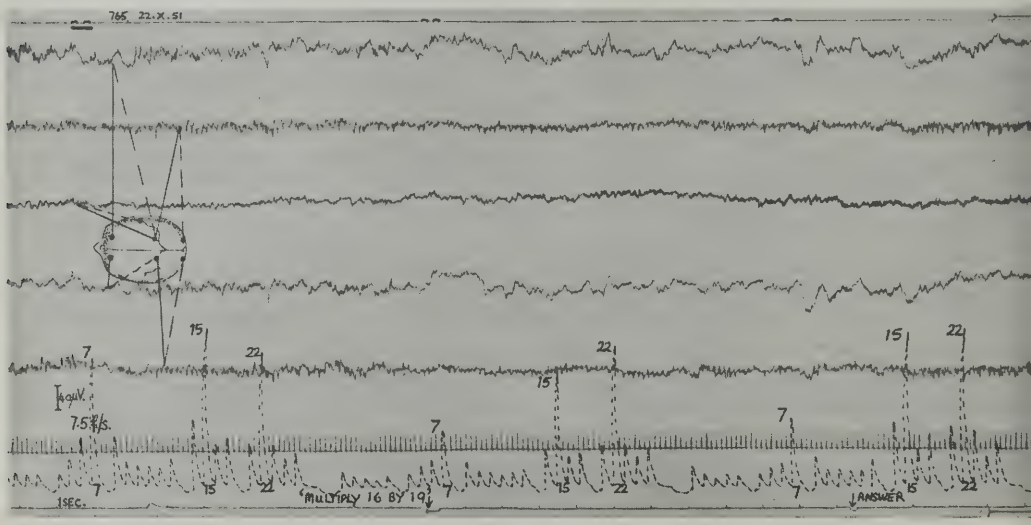


Fig. 1. Normal female, 21 yr., university student. Theta rhythm evoked during flicker and accompanied by strongly unpleasant feelings of confusion, desire to escape since it was unbearable, and profuse sweating. Activity at 6 c./s. is evoked at 12 f./s. (2nd subharmonic), but as stimulation continues and the emotion increases a non-harmonically related 5 c./s. component appears and partially outweighs the 6 c./s. rhythm. There was no prominence of theta rhythm in the unstimulated EEG. Stimulation was stopped shortly afterwards since *S* could stand it no longer. Note that the fundamental and 2nd harmonic (12 + 24 c./s.) remain unchanged despite eye opening.

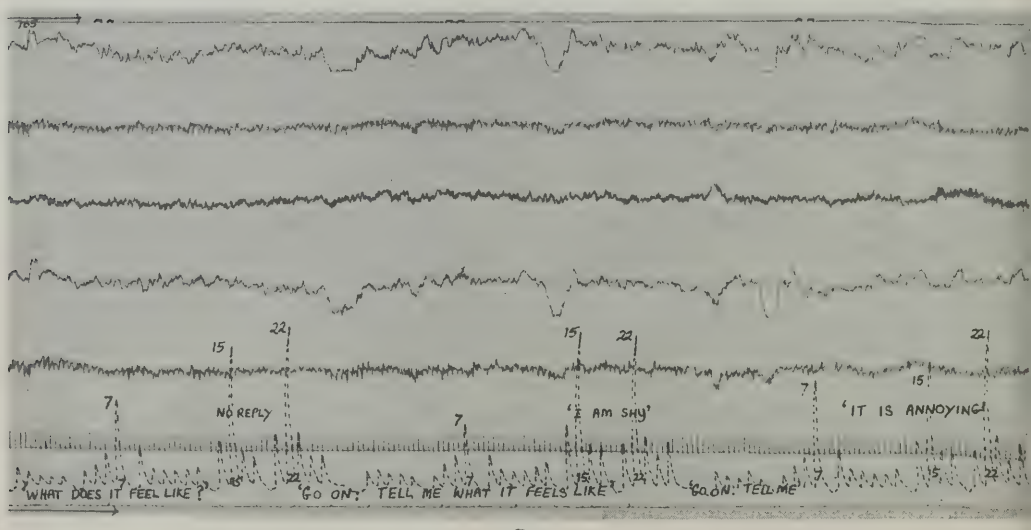
1000) capable of doing so has reported perception of visual movement, pattern and colour during stimulation within the 4–25 f./s. range. But non-visual sensations are not infrequent, particularly kinesthetic, e.g. swaying, swinging, spinning, rolling, falling and pressure, these occasionally accompanied by visceral disturbances and even vomiting. Auditory, gustatory and tactile experiences are rare. These sensory changes are usually associated with evoked or augmented activity in those cortical areas presumed to subserve their projection. All may be accompanied and augmented by emotional changes.

(v) *Intellectual*

Analysis of the sensory aspects of the flicker response grades inevitably into the intellectual, since there is no such thing as pure sensation distinct from perception: the two are blended as one and perception itself might be described as an intellectual process. But the classification 'intellectual' is necessary to differentiate response changes which occur during voluntary mental or perceptual activity, e.g. mental arithmetic, imagery,



A



B

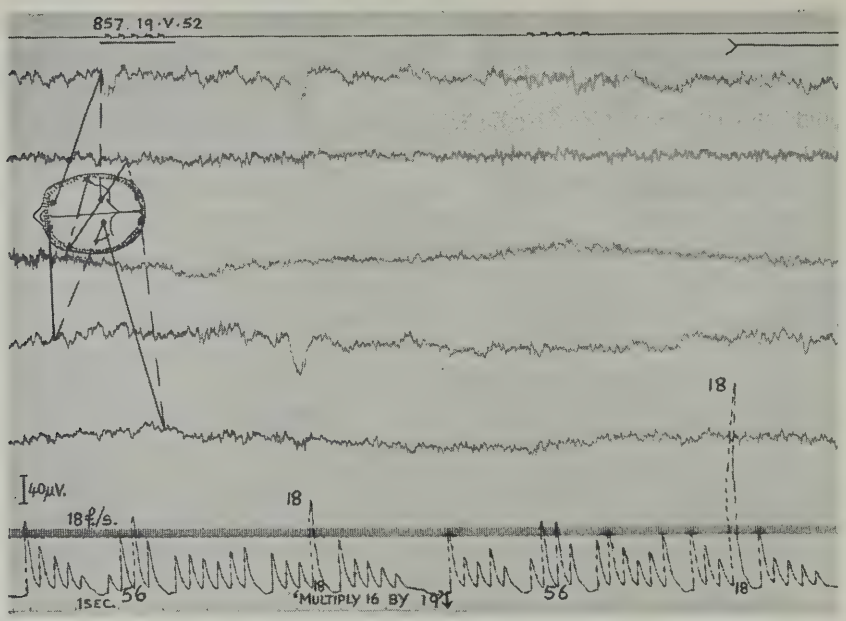
Fig. 2. A, B, normal female, 17 yr., university student. Flashes at 7·5 f./s. evoke components at 7, 15 and 22 c. s. (1st, 2nd and 3rd harmonics). During mental arithmetic all three are differentially attenuated, whilst during introspection only the 7 c./s. and to a lesser extent the 15 c./s. components are suppressed. The 7 c. s. component is identified with feelings of annoyance which are probably diminished during the intellectual activity. Feelings of no escape were also reported (samples A and B continuous.)

visual observation, etc. The principle of resonant excitation is again invoked, the assumption being that rhythmic cortical activity can be amplified either as a fundamental or harmonically related component by appropriate stimulation. In the unstimulated EEG, intellectual and perceptual activity is often accompanied by suppressed alpha and augmented beta rhythm, it having been suggested that such beta rhythm may reflect accelerated cortical activity resultant on interruption and desynchronization of alpha-scanning mechanisms (Mundy-Castle, 1951*b*). A formidable obstacle in studies of spontaneous beta rhythm is its extremely low voltage (mean amplitude approximately 8 μV., Mundy-Castle, 1951*b*). But if photic stimulation is also utilized, it is possible to augment such rhythms and hence observe more clearly how they change during perceptual and other mental operations. Preliminary observations on about 100 subjects have amply validated our earlier findings that fast rhythms may be augmented, suppressed or remain unchanged during such intellectual behaviour (Figs. 1–3). The hypothesis arising from these and earlier studies of normal spontaneous beta rhythm is that beta rhythms may act either as mediators of information concerning externally or internally derived perceptual patterns, or as components incorporated in perceptive and mental analysing and registration systems, and as such displaying blocking behaviour similar to that of the alpha rhythms. Frequencies of these distinct activities are apparently different, the latter often harmonically related to alpha frequencies, the former displaying all frequencies from 14 to 30 c./s.

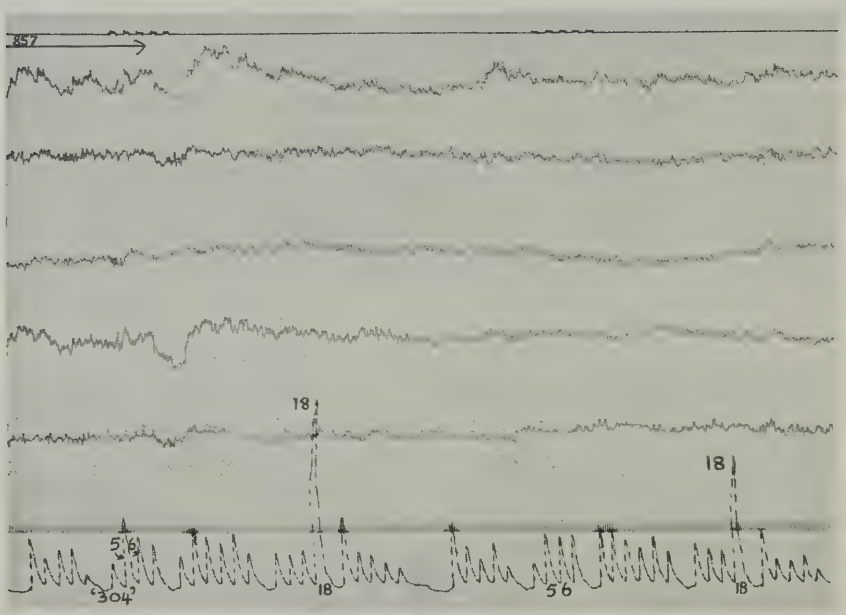
The chief difficulty in proceeding further than this lies in the excessive complexity of sorting out and comparing changes in 800 or more possible variables, despite the aid of automatic frequency analysis. The toposcopic developments of Walter & Shipton (1951) may help to solve some of these problems, since the toposcope provides simultaneous and immediate information about many of the components of a response pattern at a particular stimulus frequency. A personal communication from Dr Walter implies that evidence has already been produced to suggest a process of dissection of grouped or time-patterned stimuli in the projection areas followed by later reformation in the frontal and particularly temporal regions. But even with toposcopic display such experiments present an excessive amount of data and considerable difficulty in complex analysis is inevitable.

IV. Flicker response and cortical excitability level

The primary flicker response characteristics, which are unrelated to resonance effects, have not yet been mentioned. An observable cortical response usually occurs to a single bright flash. If a rhythmic succession of flashes is given, a train of primary responses is usually set up at the same frequency, often associated with 2nd or 3rd harmonics. These harmonics are probably due to the fact that each response consists of a train of waves, and that with certain temporal placement of each stimulus the conditions are favourable for the appearance of activity at an average of double or treble the stimulus frequency. There are, however, marked individual differences in the occurrence of such harmonics, as there are in the upper limits of flicker following, i.e. responses recurring at the stimulus frequency. Mundy-Castle (1952*a*) has recently found a relation between frequency of alpha rhythm and ratings of the Heymans–Wiersma–Biesheuvel (1929, 1932, 1949) temperament variable Secondary Function. Forty subjects were rated by a number of assessors using a paired-comparison technique, and a product moment correlation of $r = 0{\cdot}46$



A



B

Fig. 3. A, B, normal female, 21 yr., university student. Marked augmentation of evoked fundamental 18 c./s. component during mental arithmetic with flashes at 18 f./s. The 18 c./s. rhythm shows a fairly slow die-away after the response to the problem. This phenomenon was quantitatively repeatable. Note the slight increase of the non-harmonically related 5 c./s. component as opposed to the decrease of the 3rd sub-harmonic (6 c./s.) during the solution of the problem (samples A and B continuous).

($P < 0{\cdot}01$) was found between the individual mean rating values and alpha frequencies. This relation has been accepted as a working hypothesis, and is explained as due to constitutional differences in central nervous excitability characteristics. Thus the high alpha frequencies of the quick, impulsive, variable 'primary functioning' individuals are attributed to a high level of excitability, the low alpha frequencies of the slow, cautious, steady 'secondary functioning' individuals to low-level cortical excitability. Such differences in central nervous excitability have been suggested by several other studies of Secondary Function (Biesheuvel, 1938; Heymans, 1929; Wiersma, 1932). A statistical analysis of the incidence of high frequency flicker following (14–25 c./s.) and 2nd and 3rd harmonics in 154 young normal adults revealed that subjects with low alpha frequencies ($< 10{\cdot}3$ c./s.) showed a significantly smaller incidence in all these categories than did subjects with high alpha frequencies (Table 3). These findings confirm the above suggestions concerning differences in excitability characteristics. A high excitability level favours more rapid trains of impulses in response to each stimulus, thereby increasing the probability of discharge at double or treble the stimulus frequency. This is enhanced by quicker recovery-time of cortical neurons, as is the occurrence of higher frequency flicker following. It should be noted that there is no significant relation between spontaneous

Table 3. *Relations between mean alpha frequency and evoked activity*

(154 normal adults (mean age = 22 yr., S.D. 5 yr.))

	Following at 14–26 c./s.		With 2nd harmonic		With 3rd harmonic	
Evoked activity	Yes	No	*N*	%	*N*	%
Mean alpha frequency $>10{\cdot}3$ c./s., $N=80$	75	6	52	65	24	30
Mean alpha frequency $<10{\cdot}3$ c./s., $N=74$	55	18	28	35	7	9
Tests of significance	$\chi^2 = 8{\cdot}6848$, $P < 0{\cdot}01$, significant		$t = 3{\cdot}8979$, $P < 0{\cdot}01$, significant		$t = 3{\cdot}3399$, $P < 0{\cdot}01$, significant	

beta rhythm and alpha frequency, so that these findings cannot be attributed simply to resonance effects.

Independent confirmation of these conclusions has recently been advanced by Gastaut, Gastaut, Roger, Corriol & Naquet (1951), who have demonstrated by totally different methods a relation between alpha frequency and cortical excitability level, together with behavioural differences closely resembling those within the range of primary-secondary function. These differences were also described as temperamental. It must be stressed, however, that the correspondences between flicker-response characteristics and alpha frequency are far from perfect and that until more is known of the factors contributing to alpha frequency they remain only as indirect and rather obscure reflections of temperamental differences.

V. Summary and conclusions

Rhythmic photic stimulation, in conjunction with electroencephalography, is an effective method for studying the psychological significance of electrical rhythms in the brain. In addition to the evocation of epileptic seizures, flicker can elicit a wide variety of perceptual and other mental states, and the evoked rhythms which accompany these variations provide valuable information concerning the electrophysiology of mental mechanisms.

Abnormal responses are not confined to epileptics, and there is evidence of a relation between abnormal responses and behaviour deviation, tentatively attributed to a constitutional thalamic instability related to immaturity. The threshold to an abnormal response in non-epileptics sometimes varies according to the clinical state of the patient, being raised to a normal level with successful treatment. All flicker responses are susceptible to variation with changes in the state of the subject tested.

There are indications of a relation between primary flicker response characteristics, central nervous excitability and the Heymans–Wiersma–Biesheuvel temperament variable Secondary Function.

Future developments in this type of study will most probably involve toposcopic display systems in addition to conventional electroencephalography and automatic frequency analysis, since under the present system the limits of analytic resolution appear to have been reached.

References

Adrian, E. D. & Matthews, B. H. C. (1934). The Berger rhythm; potential changes from the occipital lobes in man. *Brain*, lvii, 355–85.

Biesheuvel, S. (1938). The measurement of threshold for flicker and its value as a perseveration test. *Brit. J. Psychol.* xxix, 27–38.

Biesheuvel, S. (1949). An observational technique of temperament and personality assessment. *Bull. Nat. Inst. Pers. Res.* (*S. Afr.*), i, no. 4, 9–27.

Daly, D. & Bickford, R. G. (1951). Electroencephalographic studies of identical twins with photo-epilepsy. *EEG Clin. Neurophysiol.* iii, 245–9.

Fairweather, D. S., O'Sullivan, H. J. L. & Walter, W. G. (1949). Unverrichts's myoclonic epilepsy in identical twins. *EEG Clin. Neurophysiol.* i, 115–16.

Gastaut, H. (1949). Effets des stimulations physiques sur l'E.E.G. de l'homme. *EEG Clin. Neurophysiol.*, Suppl. no. 2, pp. 69–82. Paris: 2e Congres de l'Électroencephalographie.

Gastaut, H. (1950). Combined photic and Metrazol activation of the brain. *EEG Clin. Neurophysiol.* ii, 249–61.

Gastaut, H. & Hunter, J. (1950). An experimental study of the mechanism of photic activation in idiopathic epilepsy. *EEG Clin. Neurophysiol.* ii, 263–87.

Gastaut, H., Gastaut, Y., Roger, A., Corriol, J. & Naquet, R. (1951). Étude électrographique du cycle d'excitabilité cortical. *EEG Clin. Neurophysiol.* iii, 401–28.

Heymans, G. (1929). *Inleiding tot de Speciale Psychologie*, 2 vols., 242 and 260 pp. Haarlem: De Erven, F. Bohn.

Hill, D. (1944). Cerebral dysrhythmia; its significance in aggressive behaviour. *Proc. R. Soc. Med.* xxxvii, 317–30.

Hill, D., St J. Loe, P., Theobald, J. & Waddell, M. (1951). A central homeostatic mechanism in schizophrenia. *J. Ment. Sci.* xcvii, 111–31.

Mundy-Castle, A. C. (1952*a*). The electroencephalogram in relation to temperament. Fourth Annual Congress of the South African Psychological Association, 1951. To be published.

Mundy-Castle, A. C. (1951*a*). A case in which visual hallucinations related to past experience were evoked by photic stimulation. *EEG Clin. Neurophysiol.* iii, 353–6.

Mundy-Castle, A. C. (1951*b*). Theta and beta rhythm in the electroencephalograms of normal adults. *EEG Clin. Neurophysiol.* iii, 477–86.

Mundy-Castle, A. C. (1953). An analysis of central responses to photic stimulation in normal adults. *EEG Clin. Neurophysiol.* v, 1–22.

Mundy-Castle, A. C. (1952*b*). The clinical significance of photic stimulation. *EEG Clin. Neurophysiol.*

Walter, W. G. (1946). Analytical means of studying the nature and origin of epileptic disturbances. *Res. Publ. Ass. Nerv. Ment. Dis.* xxvi, 237–51.

Walter, W. G. (1950*a*). The twenty-fourth Maudsley lecture: the functions of electrical rhythms in the brain. *J. Ment. Sci.* xcvi, 1–31.

Walter, W. G. (1950*b*). Discussion on recent advances in the EEG diagnosis of epilepsy. *Proc. R. Soc. Med.* xliv, 315–21.

Walter, W. G., Dovey, V. J. & Shipton, H. W. (1946). Analysis of the electrical response of the human cortex to photic stimulation. *Nature, Lond.*, clviii, 540–1.

Walter, W. G. & Shipton, H. W. (1951). A new toposcopic display system. *EEG Clin. Neurophysiol.* iii, 281–92.

Walter, V. J. & Walter, W. G. (1949*a*). The central effects of rhythmic sensory stimulation. *EEG Clin. Neurophysiol.* i, 57–86.

Walter, V. J. & Walter, W. G. (1949*b*). The effects of physical stimuli on the EEG. *EEG Clin. Neurophysiol.*, Suppl. no. 2, pp. 60–8. Paris: 2e Congres de l'Électroencephalographie.

Wiersma, E. D. (1932). *Lectures on Psychiatry*, pp. 610. London: H. K. Lewis and Co.

(*Manuscript received* 8 *November* 1952)

THE DYNAMIC FORCES IN THE PERSONALITIES OF JUVENILE DELINQUENTS IN THE EGYPTIAN ENVIRONMENT

By MUKHTAR HAMZA

Ibrahim University, Cairo

I. *The problem* (p. 330). II. *The cases and the tests* (pp. 330–331). III. *The control group* (p. 331). IV. *Th material for analysis* (p. 331). V. *The variables* (pp. 331–332). VI. *Scoring and calculations* (p. 332 VII. *The results* (pp. 332–334); *Discussion of the results* (pp. 332–334). VIII. *Some comments on the storie* (p. 334). IX. *Study of the families* (pp. 335–337): (i) *The broad picture* (pp. 335–336); (ii) *The relevance intelligence* (p. 336); (iii) *Conclusion* (p. 337). X. *Summary and conclusions* (p. 337). *References* (p. 338).

I. The problem

The problem of juvenile delinquency is getting more serious in Cairo. Many articles hav been written on this problem from various points of view; but an objective study wa desirable, and it was felt that a psychological approach to understanding the problem wa needed. It was hoped that such a study might throw some light on the drives, needs attitudes, family relationships and some of the important environmental condition influencing delinquency.

Among various techniques considered for carrying out the investigation was the inter view. The investigator began by interviewing a number of juvenile delinquents, but h soon discovered the great difficulty in getting them to talk freely about their problems An indirect approach was therefore necessary. It was decided to apply one of the projec tion tests. The Thematic Apperception Test (T.A.T.) was chosen for this purpose.

Neither this test nor any of the personality tests are in use in Egypt. This fact in itsel constituted a problem which the investigator had to overcome at the start. There are tw problems, therefore, in the investigation which this article reports:

(1) The standardization of the T.A.T. test, and its application to a control group and a experimental group of juvenile delinquents.

(2) The analysis of the stories in order to study the dynamic forces which affect th juvenile delinquents.

II. The cases and the tests

Eighty boys, whose age-range was 8–12 years, were chosen for this purpose. They wer all living in an institution which belongs to the Egyptian Union for Child Welfare, wher it is hoped they will be adjusted in time.

Before joining this institution, they were taken into custody by the police for variou charges. The juvenile court found it better for them to change their environment and no to live with their parents, if there were any. In the institution, they lived a family lif with a foster mother for every twenty children. A number of social workers and specialist including doctors and educationists, took care of these boys. Files and records wer always kept up to date showing all the important facts about every child. These i addition to a complete case study for each child, undertaken by the social case worke

under the guidance of the investigator, provided the original data for this study. The results of the various psychological tests were recorded. The tests included:

(*a*) The Stanford-Binet Test. (Kabbany's Arabic Revision.)

(*b*) Drever and Collins Performance Test.

(*c*) Porteus Maze Test.

(*d*) Alexander's Passalong Test.

Then, the T.A.T. test was applied and interpreted in the light of the other test results and the case studies.

III. THE CONTROL GROUP

It was found necessary to include in this study a control group of a reasonable size and of the same age-range. We therefore selected eighty boys for this purpose from some of the evening clubs situated in the same district. They belonged to families of the same standard as that of the experimental group. Other points were taken into consideration before finally deciding on this group as a suitable control group for this study.

IV. THE MATERIAL FOR ANALYSIS

Twenty cards from the T.A.T. test were given to every child in two sittings. So we had 3200 stories to analyse. Results of the analysis were grouped under two main categories:

(*a*) The force or forces emanating from the hero of the story.

(*b*) The force or forces emanating from the environment, or, as sometimes called, the 'Press'.

These are discussed in the following section.

V. THE VARIABLES

The forces or, as we shall call them, the variables reflected in the material were numerous. They differed from one case to another. We could take into account only those which were common to the majority of cases. This limited our list of variables. It was finally decided to include in category (*a*) ten variables. These were:

(1) *Achievement*: meaning persistence in doing work; the creating of something; the leading of a group; striving to accomplish something creditable; or showing signs of ambition.

(2) *Aggression*: this is regarded as 'a vital *internal* instinctive drive clamouring incessantly for repeated expression—expressions which in themselves must be pleasurable and are thus difficult to modify or repress' (Gardner, 1948). Aggression takes four different forms: (*a*) emotional and verbal; (*b*) physical, social; (*c*) physical, non-social; (*d*) destruction.

(3) *Dominance*: to restrain; imprison; to try to influence the behaviour or ideas of others.

(4) *Intraggression*: to suffer from a sense of guilt or from feelings of inferiority; or to punish oneself physically, etc.

The rest of the ten variables are abasement, nurturance, passivity, sex, succourance, and intranurturance or self-pity. These expressions are all used in the same sense as that described by Murray (1938, 1943). In addition to these ten drives, needs and attitudes, we included in our list the three following inner states and emotions: conflict, emotional change, and dejection.

In category (*b*) eight variables were included. Some of the expressions were used in category (*a*) but with different meaning. For example, 'verbal aggression' in category

(*a*) means that the hero criticizes, belittles, and threatens others, but here it means that he is criticized, belittled and threatened by others. Our eight variables are: affiliation, aggression, dominance, nurturance, rejection, lack, loss and exposure to dangers. Their meanings will be clear in § VII.

VI. Scoring and calculations

We followed the standard method of scoring (Murray, 1943; Wyatt, 1947). Each of the forces was given a score from 0 to 5 according to its strength. The criteria of strength are intensity, duration, frequency and importance in the plot.

The average of each of the forces or variables is calculated for each child and then for each group. The significance of the differences between the means for the two groups is examined by the well-known '*t*' test. In our case where we have 79 degrees of freedom, '*t*' is significant if it is equal to or more than 1·991.

The results of the two groups are tabulated to facilitate comparison.

VII. The results

For each group of variables, we constructed a table: Table 1 for the hero's needs, attitudes and emotions (category '*a*'), and Table 2 for the press or environmental forces (category '*b*'). In each table, we included the scores for the juvenile delinquents and for the control group.

For each variable, we showed the average score for the group and the range of scores as shown below.

Discussion of the results

Table 1

It is now clear that the strengths[1] of the various forces are different for the two groups. This is specially interesting in the case of those variables in which the differences between the means (or averages) are statistically significant (as shown by the '*t*' test) namely:

(1) *Non-social physical aggression*: where the hero attacks another person or kills him unlawfully; or when he fights against legally constituted authorities, or when he avenges an injury with excessive brutality.

In the case of juvenile delinquents, we were puzzled to note the enormous patience and the great painstaking expenditure of energy involved in the formulation, planning and execution of an act of aggression.

(2) *Destruction*: where the hero burns or destroys a physical object, or when he attacks or kills an animal. Usually the destroyed objects belong to persons whom he envies or hates for different reasons.

The act of destruction is in itself a pleasurable output of energy. The preparation and planning for a destructive act is important in itself, and may even be a substitute for the actual value of the thing to be stolen, or the pleasure involved in actually depriving another of his ownership.

(3) *Intragression*: in some cases, the hero blames himself for such behaviour as mentioned before, and his punishment to himself might extend to committing suicide.

(4) *Abasement*: in other cases, the hero apologizes and confesses that his behaviour was wrong. He suffers sometimes from a disagreeable situation: injury, insult or defeat. He lets it pass just to live and to make life easier for his family and himself.

[1] See the criteria of strength in § VI.

(5) *Sex*: if the hero feels lonely in solitude, he seeks the company of the opposite sex. Sometimes he gets married.

In most cases of marriage, the juvenile delinquents think of the illegal marriage, and they talk about the bastard children, their adoption and all the rest of it.

(6) *Succourance*: some heroes show their great need of support, encouragement, care and protection of someone else. The hero enjoys receiving sympathy. 'But, it is seldom one receives it.'

Table 1

	Juvenile delinquents		Normals		
The forces	Average (M_1)	Range	Average (M_2)	Range	$M_1 - M_2$
1. Achievement	12	2–35	17	8–43	−5
2. Aggression					
(*a*) Emotional and verbal	14	6–36	12	5–30	2
(*b*) Physical, social	12	4–28	8	1–23	4
(*c*) Physical, non-social	23	9–45	8	3–21	15*
(*d*) Destruction	17	5–32	6	4–18	11*
3. Dominance	16	8–41	14	1–32	2
4. Intraggression	18	12–35	9	1–18	9*
5. Abasement	21	2–33	10	6–22	11*
6. Nurturance	26	18–35	12	5–32	14
7. Passivity	12	2–51	14	4–43	−2
8. Sex	24	5–35	11	0–25	13*
9. Succourance	16	5–42	9	6–20	7*
10. Intranurturance	15	6–36	12	4–23	3
11. Conflict	19	3–46	13	5–22	6*
12. Emotional change	25	1–52	16	1–30	9*
13. Dejection	29	11–45	15	3–40	14*

* The difference between the means for the normals and delinquents passed the '*t*' test of significance.

Table 2

	Juvenile delinquents		Normals		
The press	Average M_1	Range	Average M_2	Range	$M_1 - M_2$
1. Affiliation	5	0–18	11	1–22	−6
2. Aggression					
(*a*) Emotional and verbal	12	5–27	8	0–16	4
(*b*) Physical, social	8	1–19	5	2–12	3
(*c*) Physical, non-social	13	2–25	4	0–9	9*
(*d*) Destruction	11	2–30	5	3–14	6
3. Dominance	36	18–65	25	12–40	11*
4. Nurturance	16	5–24	18	2–28	−2
5. Rejection	19	10–42	6	1–17	13*
6. Lack	26	12–53	5	0–19	21*
7. Loss	21	9–32	9	2–23	12*
8. Exposure to dangers	27	16–42	7	4–18	20*

* See Table 1 for meaning.

(7) *Conflict*: in particular, they mention the moral conflicts between the hero's needs, desires and the impulses.

(8) *Emotional change*: to the thinking of the juvenile delinquents, the heroes are unstable in most cases. They are unreliable, unpredictable and inconsistent for they fluctuate very rapidly in their mood or temper. They also like to change their jobs, to change their places of living, and, worse than all, to change their wives and children.

(9) *Dejection*: but, in their opinion, such people are not happy in their new lives. 'On the contrary, they are usually disappointed, and they remain depressed and unhappy for the rest of their lives.'

Table 2

Here again, we find statistically significant differences in the strength of some of the environmental forces affecting the two groups. This holds true of the following:

(1) *The non-social physical aggression*: the heroes of the stories of the juvenile delinquents find it difficult to live. 'It seems that everybody is against them: policemen, parents and even companions of their own age.'

(2) *Dominance*: in the case of the delinquents, the hero 'has to obey orders wherever he goes. He is always exposed to commands which are in many cases inconsistent. But still, he has to obey!'

(3) *Rejection*: 'and yet, people are indifferent to him, and they refuse to help him even in carrying out their orders. The poor fellow is unpopular and he is continuously fired from all jobs.'

(4) *Lacks*: 'accordingly, he finds no opportunities and he remains poor and unhappy. No friends, no jobs, no money, and no sympathy!'

(5) *Loss*: the hero 'must live and must get money somehow: to sell his furniture, to sell his clothes, until one day he sells all his possessions. Every member of the family has to do the same. He must try hard all day and until late at night, even the girls in the family.'

(6) *Exposure to dangers*: 'this naturally exposes all members of the family to *all* sorts of dangers and injuries: physical and *social*. Worse than all is the dispersion of the family all over the place until one day, there will be no family!'

VIII. Some comments on the stories

Since the T.A.T. test is not meant for Egyptians, it was found that two or three cards were quite unsuitable. The cards which proved more stimulating were the following: 3BM, 4, 7BM, 9BM, 13, 17BM, 20. Here are some of the observations on the stories given by each of the two groups of boys:

(1) In the case of the control group, about 45% of the stories (nine out of twenty) were psychologically irrelevant, composed, for the most part of inventions of the moment. None of these are representative of a determining tendency in the personality of the subject.

(2) The percentage, in the case of the juvenile delinquents, goes down to 30% (six out of twenty).

(3) Contrary to expectation (Jackson, 1951), in the case of card no. 16, 'The Blank Card', the delinquents did not bother about it, while the subjects of the control group closed their eyes and invented long stories.

(4) In all possible ways, the juvenile delinquents expressed their desire for a stable life, their hope of learning suitable trades and their dream of living in *homes* of their own, quietly and happily.

(5) In general, the endings of the stories of the delinquents were in most cases unhappy and full of grief: prison, death, fire, committing suicide and the like.

(6) The boys express their sorrow for such unhappy endings and for the unsatisfactory behaviour leading to such situations.

(7) In some of the stories of the delinquents, they talked about two heroes in the same story; which means that two forces of the subject's personality are represented by two different characters, i.e. 'endopsychic thema'.

IX. STUDY OF THE FAMILIES

(i) *The broad picture*

In studying the families of both groups, we noticed that for the control group there were relatively few broken homes compared with the incidence for the delinquent group. We also noticed that the broken homes appeared not to belong to any special economic category. This was not the case in the experimental group. Here, the broken homes were usually in the lowest standards from the economic point of view. In order to check this impression statistically arbitrary ratings on a five-point scale were made of the home conditions and the economic standard of all the families. We called the five categories in each case A, B, C, D and E. The basis of classification was as follows:

Home conditions

A. Parents live together happily with all their children except for the one child in the institution. Members of the family co-operate as one unit. The family is integrated and family gatherings take place very frequently and they are enjoyed by everybody.

B. Relation between father and mother is nearly at its best, but one or two members of the family are not keen on family gatherings.

C. This is the usual standard of relation, that found in most cases. The family has its own house, flat or rooms. It is the centre for every member of the family. Father and mother must spend the nights at home. Gatherings are held occasionally.

D. There is a centre for the family, but it does not attract its members. One of the parents does not live there. In some cases, some of the children too live outside.

E. No centre, no integration, a wide dispersion of family members. Each member might live anywhere, not knowing what is happening to any other member. He is not even interested.

Table 3. *The control group*

Economic standard	Home conditions A	B	C	D	E	Total
A	3	4	4	1	0	12
B	4	5	6	2	1	18
C	6	8	11	3	2	30
D	1	1	4	4	3	13
E	1	1	1	2	2	7
Total	15	19	26	12	8	80

Table 4. *The experimental group*

Economic standard	Home conditions A	B	C	D	E	Total
A	3	4	1	1	0	9
B	1	5	7	2	1	16
C	0	3	7	7	8	25
D	0	0	1	4	9	14
E	0	0	1	5	10	16
Total	4	12	17	19	28	80

The economic standard[1] (*monthly income*)

A, more than £E 25; B, 20–25; C, 10–20; D, 5–10; E, less than 5.

The distribution of each group amongst these different categories will now be shown in the form of contingency tables (see Tables 3 and 4).

[1] This scale is only true of the poor classes included in this study.

The most striking feature in Tables 3 and 4 is the distribution of the lower ratings Take, for example, the horizontal classes D and E in the second table. None of the four teen or the sixteen poor families belongs to the well-integrated families as shown by the vertical columns A and B. This, and other facts obvious from the tables, suggest that there was a strong relationship between poverty and home conditions. As the expected frequency in each cell was sometimes less than 5 we could not apply χ^2 to the data as they stood. We therefore grouped the frequencies into three classes only α, β and γ where $\alpha = A + B$; $\gamma = D + E$ and $\beta = C$. Then we calculated χ^2. The results for Table 3 were $\chi^2 = 14{\cdot}1958$ and for Table 4 $\chi^2 = 38{\cdot}3334$.

For four degrees of freedom χ^2 is significant at $P = 0{\cdot}01$ if $\chi^2 = 13{\cdot}277$. These results are therefore highly significant. We can, thus, conclude that the poverty and the broken homes were related to each other.

(ii) *The relevance of intelligence*

In spite of this general conclusion, we find, in Table 3, some poor families which are well integrated and whose homes are not broken. These seem to be the more intelligent people (as judged by the results of their children). We classified the I.Q.'s into five levels as follows: A, I.Q. greater than 110; B, 100–110; C, 90–100; D, 70–90; E, less than 70.

Then, we constructed contingency tables (Tables 5 and 6) as before for the two groups

Table 5. *The control group*

Economic standard	Home conditions A	B	C	D	E	Total
A	3	2	1	0	0	6
B	4	5	6	1	0	16
C	7	11	11	5	3	37
D	1	1	7	5	4	18
E	0	0	1	1	1	3
Total	15	19	26	12	8	80

Table 6. *The experimental group*

Economic standard	Home conditions A	B	C	D	E	Total
A	2	3	1	0	0	6
B	1	6	9	2	1	19
C	1	2	4	9	10	26
D	0	1	2	4	8	15
E	0	0	1	4	9	14
Total	4	12	17	19	28	80

Here again, the figures speak for themselves, and the relation is obvious. Grouping A and B, D and E and calculating as before, we found that for Table 5, $\chi^2 = 17{\cdot}7743$; and for Table 6, $\chi^2 = 42{\cdot}875$.

In both cases the results are significant. There appears therefore to have been a close relationship between 'home condition' and the 'level of intelligence'. This interesting finding is in agreement with Burt's studies in the L.C.C.

(iii) *Conclusion*

We may therefore conclude that 'poverty' and 'poor intelligence' were two major factors contributing to broken homes. This in itself is one of the most important results bearing on the problem of juvenile delinquency.

X. SUMMARY AND CONCLUSIONS

It was evident from the case-studies and the testing programme that in every case of delinquency there were several factors contributing to the problem. This finding tallies with those of other researches in the field and with Burt's (1924) conclusion, that: 'The nature of these factors, and of their varying combinations, differs greatly from one individual to another; and juvenile offenders, as is amply clear, are far from constituting a homogeneous class.'

However, we can group these factors under two main headings:

Factors within the child

These include his needs, desires, interests, attitudes and emotions. Examples of these are his need for care and protection, his need for security in life and his need for guidance and education. We mean education which is suitable for the child's abilities. In this study, we found that most of the delinquents were below the average intelligence. The mean I.Q. for the whole group was no more than 73 which shows clearly that their abilities were not promising. At least 25% of the boys were mentally deficient.

Factors from without (environmental forces)

These include *the home conditions* and the relations between the parents. Most of our cases belonged to broken homes: one or both parents dead, deserted, in prison, divorced, etc. Very much related to this point is *poverty*. Its evil results are well known.

Bad home conditions and poverty lead to all sorts of complications. This also means bad conditions *outside* the home: no control, poor jobs, bad companions, direct and indirect influences of adult companions, cinemas, and other places where offences are committed.

Knowing so much about the hard conditions and the disadvantages the delinquent suffers, are we to blame him? Is he completely and thoroughly responsible for his acts?

This research gives a negative answer. The findings indicate that the efforts of the social reformers should be directed towards the environmental forces more than anything else, and the problem will remain unsolved until the fundamental causal factors in the society itself are overcome.

The writer hopes to have shown that the dynamic forces in the juvenile delinquents can be studied objectively by means of projection tests. It was found here that some of the cards of the T.A.T. test are inadequate for the Egyptians, others could be improved upon. Once this modification is completed; it is hoped to apply the test on a large scale and to publish norms for both sexes and for different ages.

References

Burt, C. (1924). *The Young Delinquent.* London: University of London Press.
Gardner, G. E. (1948). 'The Community and the Aggressive Child.' Paper read in the International Conference on Child Psychiatry, London.
Jackson, Lydia (1951). Emotional attitudes towards the family of normal, neurotic and delinquent children. *Brit. J. Psychol.* xli.
Murray, H. A. & Morgan, C. D. (1938). *Explorations in Personality.* New York.
Murray, H. A. (1943). *Thematic Apperception Test Manual.* Cambridge, Mass.
Wyatt, F. (1947). *The Scoring and Analysis of Thematic Apperception Test.* New York.

(*Manuscript received* 19 *February* 1953.)

THE EFFICIENCY OF DIFFERENTLY CONSTITUTED GROUPS OF CHILDREN IN DIFFERENT TYPES OF TASKS

By F. V. SMITH, W. SLUCKIN and D. GRAHAM

Department of Psychology, University of Durham

I. *Introduction* (p. 339). II. *Preliminary work* (p. 339). III. *Experimental procedure* (pp. 339–342). IV. *The structure of the groups* (pp. 342–343). V. *Results* (pp. 343–345): (i) *Ratings of observers* (p. 343); (ii) *Analysis of relations between variables* (pp. 343–344); (iii) *Contribution of group members* (pp. 344–345); (iv) *Group cohesion* (p. 345); (v) *Correlations established* (p. 345). VI. *Conclusions* (pp. 345–346). *Reference* (p. 346).

I. Introduction

Many of the reasons why groups vary in the ease, efficiency and enjoyment with which tasks are accomplished are undoubtedly of psychological significance. It is conceivable, too, that the best composition of the group in terms of happiness and efficiency will vary for different kinds of tasks. The present paper describes a procedure designed to obtain empirical data about these possibilities.

II. Preliminary work

The intention was to design three different types of task which would be attractive to and capable of sustaining the interest of the age groups of children likely to be available for the experiment. Over a period of 8 weeks and with twenty children, a variety of tasks were tried. Attention was directed to the enjoyment shown by the children, the size of the group, the ease of manipulation, the suitability of the task for scoring, the nature of the instructions and the times to be allowed for each task. The tasks described in § III were evolved. All phases of the experiment were carried out in the observation room within the Department, which enabled the observers to see and hear everything that the children did without the children being aware of observation.

III. Experimental procedure

Task I. Building a castle. The group was provided with sixty-four large, smoothly finished oak blocks—see photograph. One experimenter (the same person throughout the experiment) brought the children into the room. After allowing a few minutes for the children to become accustomed to the new surroundings and ask questions about the fittings in the room, the experimenter began—'Now those games in there (Binet material in display cabinet) are for younger children. We have here some blocks and planks for older people like you.' The children examined the materials. 'Now I want you all to work together and build a fine castle here on the mat (in the centre of the room). I shall now go away and leave you to build it yourselves.' The time allowed was $7\frac{1}{2}$ min. The experimenter then left the room and joined the two other observers in another room from which the children could be watched. The observations were recorded independently by the three observers on blanks of the kind illustrated on the following page.

OBSERVATION BLANK

Task no.: Observer: Date:

Age group: Group code:

	Quantity:						
Quality							
	1	2	3	4	5	6	7
Friendliness							
	1	2	3	4	5	6	7
Co-operation							
	1	2	3	4	5	6	7

	Complete (C)	Incomplete (I)
Group cohesion		

	Uniform (U)	Non-uniform (N)
Contribution of group members		

Remarks:

The ratings of Friendliness, Co-operation and Quality (tasks I and III) were made with reference to the scales defined on p. 341.

During the preliminary phase, it was apparent that friendly behaviour between the members of the group did not necessarily involve co-operation. The ratings under 'co-operation' applied to positive acts of co-operation, such as the passing of blocks between members or the holding of some blocks in position while another member built a supporting column. 'Group cohesion' was recorded as complete when all members continued to be interested in the project for the entire period, and 'incomplete' if one or more children wandered away from the centre of activity. The contribution of group members was recorded as 'non-uniform' if one or more children were outstanding in their contribution. When all three observers could not agree, the rating was 'doubtful'. 'Quantity' referred to the number of blocks actually forming part of the castle.

When the time had elapsed, the experimenter entered the room and having admired the castle and asked a few questions about the design, requested the children to stack the blocks in the corner to make ready for the next task.

Task II. Routine classification at speed. This task involved 118 smaller blocks in nine different shapes (see Pl. 1). These were heaped in random order on a large tray. The experimenter began, 'Here are a number of blocks of different shapes and sizes all mixed up together. I want you all to sort them out and place all blocks of the same kind together on a separate sheet of paper along the edge of the carpet—like this (a demonstration was given with four blocks). Now I shall go away again. Are you ready? Go!' The time allowed was $2\frac{1}{2}$ min. The observations recorded were the same as for task I except that the number of blocks correctly classified was recorded under 'Quantity' and no estimate of 'Quality' was recorded.

Task III. The building of the dolls' house. (Construction and thoughtful arrangement with ample time.) The experimenter began, 'Here we have some furniture for a house and the people who live in the house' (see Pl. 1). The children invariably examined the items with considerable excitement. 'Now (taking up some planks) here is *one* room of the house.'

Rating scales used

Friendliness

1	2	3	4	5	6	7
Fighting, snatching, overt aggression	Snatching, spiteful undoing of the work of others; unfriendly argument but no overt aggression	Some unfriendly undoing of the work of others but no snatching	No bias in either direction	Some overt indications of friendliness; no indications of unfriendliness whatsoever	Consistently, words, smiles and gestures of friendliness	As for 6, with the addition of acts of affection, e.g. embraces

Co-operation

Group working *entirely* as individuals	Some recognition of others, desultory discussion but no positive acts of assistance	Some discussion and a few isolated acts of assistance without reference to an agreed design, e.g. passing a block	Some consultation on design; some tacit allocation of functions; relevant mutual assistance	Consultation on design and efficient allocation of functions but not extending to many phases of the task	Co-operation, efficient division of labour extending to most phases of the task	Co-operation involving planned division of labour which extended over the whole period of the task

Quality

The quality of the 'Castle' (task I) and of the arrangement of the dolls' house (task III) was assessed on a seven-point scale. In task I (castle) the integration of the design, the stability of the structure and the functional value of the parts influenced the ratings. With task III (the dolls' house) the number of rooms, the appropriate placing of furniture, doorways and persons (grandfather, mother, baby, maid) influenced the ratings. In both tasks, experience during the preliminary phase had indicated what could be expected from children of this age.

Planks were arranged on the floor allowing space for a doorway. 'I want you to build a house and put the furniture and the people in it in the best way. I shall go away now and leave you to finish it.' The time allowed was 7 min. The observations recorded were the same as for task I, except that the number of planks used was not recorded.

For all three tasks, each observer used a separate observation blank and made his ratings independently of the others.

IV. The structure of the groups

The groups consisted exclusively of girls. The age was as near 10 as possible within the resources of the school, with an average age on the day of the experiment of 9·8 years, $\sigma = 12{\cdot}3$ months. The school was a small one, situated near the Department. The Headmistress knew all the children very well. On consideration, it was thought that she was better equipped than the individual class teachers to rate all the children. The children were rated by the Headmistress as being of high, medium or low assertiveness. This estimate of assertiveness is of course based upon assertiveness as apparent in the activities of the school. Ability was estimated by the Otis Mental Ability Alpha Test. Altogether 88 children took part in the main experiment. They were arranged in groups of four. There were twenty-two groups, as follows:

1, 2, 3	Uniform high ability—uniform high assertiveness.
4	Uniform high ability—uniform medium assertiveness.
5, 6	Uniform high ability—mixed assertiveness.
7	Uniform medium ability—uniform high assertiveness.
8, 9	Uniform medium ability—uniform medium assertiveness.
10	Uniform medium ability—uniform low assertiveness.
11	Uniform medium ability—mixed assertiveness.
12	Uniform low ability—uniform high assertiveness.
13	Uniform low ability—uniform medium assertiveness.
14	Mixed ability—uniform high assertiveness.
15, 16, 17	Mixed ability—uniform low assertiveness.
18, 19, 20, 21, 22	Mixed ability—mixed assertiveness.

The designations were as follows:

(*a*) Assertiveness: for 'mixed assertiveness', the group consisted of one child rated as highly assertive, one rated as of 'low' assertiveness and the other two of 'medium' assertiveness. 'Uniform assertiveness' meant that all the children were in the same category, i.e. high, medium or low.

(*b*) Ability: a group was described as of 'mixed ability' when the difference between the highest and lowest I.Q. was more than 18 points. The average difference for 'mixed' groups was in fact 23·1 with $\sigma = 5{\cdot}15$.[1] It was always arranged that a 'mixed' group included one child well above the average in ability and one child well below the average. A group was described as of 'uniform ability' when the difference between the highest and lowest I.Q. was less than 18 points. In fact, the mean difference was 11·3, $\sigma = 3{\cdot}81$. The 'uniform' groups were further differentiated into 'uniform high' (all above I.Q. = 100), 'uniform low' (all below 100 I.Q.) and 'uniform medium' (comprising groups which did not fall into either of these two classes).

It was hoped to have more groups in some of the above categories; but the resources of the school could not provide them.

V. Results

(i) *Ratings of observers*

Before analysing the experimental results, the agreement between the ratings of the three observers was investigated. Seven-point scales for quality of work, friendliness and co-operation, as defined earlier, were used. The correlations between the judgements of the three raters were: 0·687, 0·682 and 0·777. In each case $r \geqslant 3\sigma r$. The conclusion that the ratings are combinable was checked by analysis of variance between the raters:

Source of variation	Sum of squares	D.F.	Variance
Between raters	4	2	2·000
Within raters (error variance)	696	173	4·023

F was not significant, thus confirming that the combination of the ratings was justified.

(ii) *Analysis of relations between variables*

Several analyses of variance were then carried out to determine the relationships between each of the two independent and the various dependent variables.

(*a*) *Composition of ability and friendliness*

Source of variation	Sum of squares	D.F.	Variance
Between	105	2	52·5
Within	388	63	6·159

$F = 8{\cdot}52$, significant at the 1 % level. Groups of high ability were more friendly (average rating 5·0) than groups of low and medium ability combined (average friendliness rating 3·9; $t = 4{\cdot}86$, significant at 1 % level) and more friendly than groups of mixed ability (average friendliness rating 4·3; $t = 2{\cdot}81$, significant at 1 % level).

(*b*) *Composition of ability and co-operation*

Source of variation	Sum of squares	D.F.	Variance
Between	91	2	45·5
Within	402	63	6·381

$F = 7{\cdot}13$, significant at the 1 % level. Groups of high ability were more co-operative (average rating 4·9) than groups of low and medium ability combined (average rating 4·2; $t = 3{\cdot}15$, significant at 1 % level) and more co-operative than groups of mixed ability (average co-operative rating 4·0; $t = 11{\cdot}32$, significant at 1 % level).

[1] This figure is influenced by two cases where the range was 35 and 28 points of I.Q.

(*c*) *Composition of ability and quality of work in task I*

Source of variation	Sum of squares	D.F.	Variance
Between	163	2	81·5
Within	252	19	13·263

$F = 6{\cdot}145$, significant at the 1% level. The quality of work of groups of 'mixed' ability was below (average rating 2·7) that of groups of high ability (average quality rating 4·8; $t = 2{\cdot}95$, significant at 5% level), and below that of groups of medium and low ability combined (average quality rating 4·4; $t = 2{\cdot}88$, significant at 5% level).

(*d*) *Composition of ability and quality of work in task III*

Source of variation	Sum of squares	D.F.	Variance
Between	93	2	46·5
Within	174	19	9·158

$F = 5{\cdot}078$, significant at the 5% level. The quality of work of groups of high ability (average quality rating 5·5) was above that of groups of mixed ability (average quality rating 3·9; $t = 3{\cdot}38$, significant at 1% level) and above that of groups of medium and low ability combined (average quality rating 4·1; $t = 2{\cdot}51$, significant at 5% level).

(*e*) *Composition of ability and quantity of materials used in task I*

F was not significant. It should be noted that the object of this task was to build a 'good castle' rather than to use the maximum number of blocks.

(*f*) *Composition of ability and quantity of materials used in task II*

$F = 4{\cdot}36$, significant at the 5% level. In this task of classification, it appears as if output were proportional to the average ability within the group. Groups of high ability were better (average number used 111) than groups of 'mixed' ability (average number used 87; $t = 2{\cdot}67$, significant at 5% level) and better than groups of medium and low ability combined (average number used 82. $t = 4{\cdot}38$, significant at 1% level).

(*g*) Analyses of variance were carried out for 'composition of assertiveness' and (*a*) friendliness, (*b*) co-operation, (*c*) quality of work in task I, (*d*) quality of work in task III, (*e*) quantity of materials used in task II. None of the F-values were significant.

(*h*) *Composition of assertiveness and quantity of materials used in task II*

Source of variation	Sum of squares	D.F.	Variance
Between	3780	3	1260
Within	6138	18	341

$F = 3{\cdot}70$, significant at the 5% level. Groups of high assertiveness were fastest and groups of low assertiveness were slowest. As the distribution of the scores in one category was skewed, conclusions from the F-value may be somewhat misleading.

(*i*) Analyses of variance between tasks for (*a*) friendliness and (*j*) co-operation were carried out. The differences between tasks were not significant.

(iii) *Contribution of group members*

With reference to the 'contribution of group members' to the total group effort, there were three categories: uniform, doubtful and non-uniform. A contingency table relating 'contribution of group members' to tasks I, II and III was drawn up and tested by the χ^2 method. χ^2 was 12·89, significant at the 5% level. It was clear that task II could not

be combined with the other two tasks. When a table omitting task II was tested, χ^2 was not significant. Since the contribution of members was found to be non-uniform for task II in only two out of the twenty-two groups, further analysis of this problem was not pursued.

With tasks I and III combined, $N=44$. With this number, an analysis for χ^2 is just permissible (Lewis & Burke, 1949). It was found that the variable 'contribution of group members' was not significantly related to (*a*) ability, (*b*) assertiveness, (*c*) quality of work, (*d*) friendliness, (*e*) co-operation.

(iv) *Group cohesion*

The intention was to classify group cohesion as (*a*) complete, (*b*) doubtful, and (*c*) incomplete. However, during the sixty-six sessions, cohesion was complete in sixty and doubtful in the remaining six cases. Consequently, no attempt was made to relate this to any other variables.

(v) *Correlations established*

The following correlations were established:

Task I

(*a*) Quantity-quality $r=0{\cdot}352$, not significant.
(*b*) Quantity-friendliness $r=0{\cdot}210$, not significant.
(*c*) Quantity-co-operation $r=0{\cdot}624$, significant at 1% level.
(*d*) Quality-friendliness $r=0{\cdot}283$, not significant.
(*e*) Quality-co-operation $r=0{\cdot}494$, significant at 1% level.
(*f*) Friendliness-co-operation $r=0{\cdot}604$, significant at 1% level.

Task II

(*a*) Quantity-friendliness $r=0{\cdot}578$, significant at 1% level.
(*b*) Quantity-co-operation $r=0{\cdot}385$, significant at 5% level.
(*c*) Friendliness-co-operation $r=0{\cdot}697$, significant at 1% level.

Task III

(*a*) Quality-friendliness $r=0{\cdot}512$, significant at 1% level.
(*b*) Quality-co-operation $r=0{\cdot}790$, significant at 1% level.
(*c*) Friendliness-co-operation $r=0{\cdot}804$, significant at 1% level.

It may be noted that on the basis of the null hypothesis of no correlation, $\sigma_r=0{\cdot}218$; thus significance at 5% level corresponds to $r=0{\cdot}427$. The statement of significance quoted for each r above was based upon the calculation of the standard error for that particular r.

VI. Conclusions

Within the experimental situations herein described, the most interesting findings are:

(1) Groups of high ability were the most friendly and groups of medium and low ability combined were the least friendly.

(2) Groups of high ability were the most co-operative; but the least co-operative were groups of 'mixed' ability.

(3) Wherever quality was assessed, groups of 'mixed' ability produced work of lowest quality; groups of high ability produced work of highest quality.

(4) The consistent finding from the present study is that the efficiency of a group, as indicated by assessments of quantity and quality of work, co-operation and friendliness, is dependent upon the ability of members of the group and not upon their assertiveness. The one exception to this generalization occurred in task II.

(5) In task II, the greatest quantity of work was accomplished by groups of high assertiveness and the least quantity by groups of low assertiveness. With regard to ability, the usual trend was maintained, namely the greatest quantity of work was accomplished by groups of high ability. In this case, groups of mixed ability were next and then groups of medium and low ability combined.

It is clear that high output in this task is associated with *both* ability and assertiveness.

The above method of direct investigation of the performance of differently constituted groups in tasks of varying structure could obviously be applied to other age groups of children. The progression with age and the association with ratings of ability, assertiveness and other aspects of personality might provide interesting findings. Different constitutions of groups could also be investigated. We report our findings at this stage because we had exhausted the resources of the particular school and the addition of cases from other schools would have involved ratings of the children which could not have been combined with those already gained.

Our thanks are due to Miss Margaret Aspinall, Headmistress of St Hild's Demonstration School, Durham, for her sustained co-operation in this project.

Reference

Lewis, D. & Burke, C. J. (1949). The use and misuse of the chi-square test. *Psychol. Bull.* xlvi, 433–89.

(*Manuscript received* 13 *February* 1953)

AN ANALYSIS OF INTELLIGENCE TEST SCORES OF STUDENTS IN A UNIVERSITY DEPARTMENT OF EDUCATION

By A. H. D. TOZER and H. J. C. LARWOOD
University of Liverpool

I. *Introduction* (p. 347). II. *Students tested* (pp. 347–348). III. *Distribution of scores* (pp. 348–349). IV. *Norms* (pp. 349–350). V. *Variations from year to year* (p. 350). VI. *Comparison of pre-war and post-war students* (pp. 350–352). VII. *Differences between men and women* (pp. 352–353). VIII. *Differences between scores according to class of degree* (pp. 353–355). IX. *Differences between scores of arts and science students* (pp. 355–356). X. *Differences between scores according to subjects of degree* (pp. 356–357). XI. *Summary* (p. 358). *References* (p. 358).

I. Introduction

Each year from 1932–3 (except for 3 years during the war and 1 year for which results are not available) intelligence tests of various types have been given to students in the University of Liverpool Department of Education as an introduction to the subject of psychological testing. To secure comparability of results from year to year one test always used has been the National Institute of Industrial Psychology's Group Test 33. Almost 900 students have now taken this test over a period of 18 years, and the present paper discusses an analysis of some of the results obtained.

Group Test 33 is a test prepared by Burt which has been extensively used in this country since 1923. It consists of five parts, viz. *Opposites*, *Analogies*, *Mixed sentences*, *Completing sentences* and *Reasoning*, with a separate time limit for each part. Factorial studies (e.g. Slater, 1940) have shown the test to be heavily loaded with factors g and v. We have recorded the marks obtained on each of the subtests, but the present report is concerned only with total score. (The odd half mark which can be gained in part V has been dropped.)

II. Students tested

The population of the Department of Education consists of students from Liverpool and graduates from other universities taking a course of professional training for the Diploma in Education. They include '4-year students' selected before commencing their degree courses and '1-year' students selected after graduation.[1] The great majority are graduates, but a few Liverpool students admitted under the 4-year scheme who have failed the degree examination are allowed, if specially recommended, to enter the course for the Certificate in Education. The numbers of men and women tested over the whole period were roughly equal, but the 2 years, 1940–1, 1941–2 (included with the pre-war group since these students had entered the university before the outbreak of hostilities) were entirely composed of women.

In some years all students were tested; more often only some of them did the test. We tried to check the representative character of our sample by classifying the students according to sex, subject taken and kind of degree, and comparing its composition with

[1] Mean scores of these two groups did not differ significantly.

the whole population of the Department in that year, and we are reasonably confident that the groups tested were fair samples of the years from which they were taken, at any rate with respect to the variables mentioned.

III. Distribution of scores

Table 1 gives the frequency distribution of scores for men and women separately, and for all students together, for the whole period of our investigation; also, since we found some significant differences between our pre-war and post-war groups (which are discussed later), for the pre-war and post-war periods separately.

Table 1. *Distribution of scores*

	Pre-war			Post-war			All years
Scores	Men	Women	Total	Men	Women	Total	Total
185–189	3	1	4	2	3	5	9
180–184	5	9	14	6	4	10	24
175–179	17	14	31	7	4	11	42
170–174	30	22	52	19	16	35	87
165–169	29	23	52	16	14	30	82
160–164	25	31	56	15	16	31	87
155–159	58	38	96	15	18	33	129
150–154	31	33	64	10	8	18	82
145–149	20	21	41	12	16	28	69
140–144	21	33	54	7	12	19	73
135–139	22	26	48	4	11	15	63
130–134	21	17	38	5	8	13	51
125–129	10	11	21	0	3	3	24
120–124	9	9	18	3	2	5	23
115–119	1	10	11	2	1	3	14
110–114	4	3	7	1	0	1	8
105–109	1	2	3	1	1	2	5
100–104	3	0	3	0	0	0	3
95–99	5	1	6	0	0	0	6
90–94	1	0	1	0	0	0	1
85–89	0	1	1	0	0	0	1
$N=$	316	305	621	125	137	262	883

The means and standard deviations of scores are set out in Table 2.

Table 2. *Means and standard deviations of scores*

	Pre-war			Post-war			All years		
	Men	Women	Total	Men	Women	Total	Men	Women	Total
N	316	305	621	125	137	262	441	442	883
M	151·85	150·65	151·26	158·02	154·75	156·31	153·60	151·92	152·76
S.D.	18·49	17·64	18·07	16·41	15·88	16·19	18·12	16·83	17·68

The form of the score distribution may be more easily seen from the frequency polygon shown in Fig. 1. The negative skewness is obvious, and secondary peaks above and below the mode which have been reported in other investigations with university students (Heim, 1947) appear here also. The piling up of scores at the higher end of the distribution with a test not designed for a highly intelligent group is similar to that reported, for example, by Wallace (1952) with Heim's A.H.4 Test. The upper peak may be due to the overlap of the distributions of two distinct groups; the lower one, it has been suggested, may represent 'students who have entered the university on other than scholastic grounds' (Foulds & Raven, 1950). This is more than likely with our group since in selecting student

for training as teachers considerable importance is attached to non-intellectual factors of personality assessed at interview.

Statistical tests for normality (Snedecor, 1940) show that the negative skewness of the distribution is highly significant. The departure from normality with regard to kurtosis, however, is not significant.

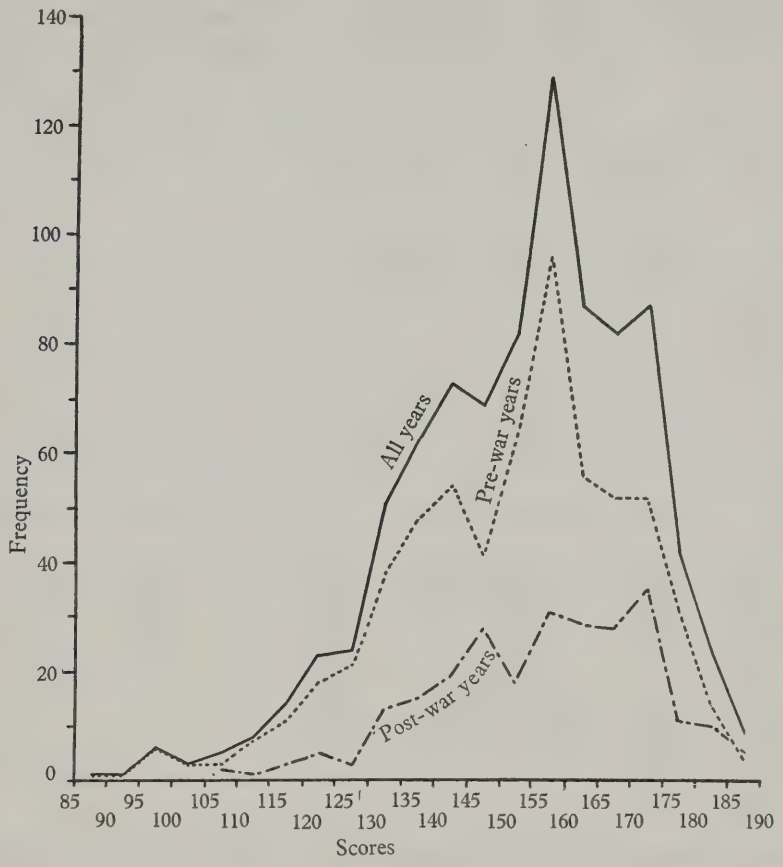


Fig. 1. Distribution of scores over all years and for the pre-war and post-war periods separately.

IV. Norms

Norms for this test are usually presented in percentile form. For comparison, therefore, we give the results for our students in Table 3. In addition to the usual percentile points we also give the scores corresponding to the 16th and 84th percentiles as an approximation to the $M \pm 1$ S.D. range. These figures may be compared with those of Vernon (1939) with a somewhat similar group of 'unsophisticated' graduates, and with the percentiles which we have calculated from the results of Babington Smith (1940) with his graduate subjects.

If we convert our raw scores to I.Q. equivalents by Babington Smith's formula which is based upon comparison between G.T. 33 and the Otis Advanced Test, and assume a S.D. of 15 points of I.Q., we find that the I.Q. corresponding to the mean score of our group is 121. From the results of 126 students who did both G.T. 33 and the Moray House Adult Intelligence Test 1 we calculate from the regression equation that our mean is equivalent

to I.Q. 126 on the latter test. This may be compared with the medians of 126 and 121 reported by Thomson (1945) for honours and pass graduates respectively (women only) in a Scottish university.

V. Variations from year to year

How constant in respect of test-intelligence has been the quality of students passing through the Department over this period of years? Table 4 shows the number of students tested each year, their mean score and the standard deviation of scores.

Table 3. *Percentile scores*

	Percentiles						
	10	16	25	50	75	84	90
Pre-war							
Men	128	133	140	155	165	170	173
Women	127	133	139	152	163	169	173
Post-war							
Men	135	142	148	160	170	173	176
Women	134	138	143	156	167	171	174
All years							
Men	130	135	143	156	167	171	174
Women	130	135	140	154	164	170	173
Vernon	128½	—	145½	158	165½	—	174
Smith	123	131	137	149	160	165	169

Table 4. *Means and standard deviations of scores of students for each year*

	Pre-war								
	1932–3	1933–4	1934–5	1936–7	1937–8	1938–9	1939–40	1940–1	1941–2
N	97	72	139	80	61	59	55	27	31
M	153·31	151·40	148·22	151·54	155·84	153·03	151·47	148·74	146·90
S.D.	16·56	16·48	17·84	19·54	17·99	19·69	20·48	14·60	17·14

	Post-war				
	1945–6	1946–7	1947–8	1948–9	1949–50
N	22	23	53	67	97
M	152·00	151·04	157·17	155·98	158·29
S.D.	18·04	17·72	12·67	15·56	17·33

Application of Bartlett's Test for homogeneity of variance shows that in spite of one suspicious looking year (1947–8) we may take it that the variance is, in fact, homogeneous ($\chi^2 = 20{\cdot}20$; D.F. $= 13$; $0{\cdot}05 < P < 0{\cdot}10$).

Examining the year-to-year variation of mean scores by analysis of variance we find the variance ratio is highly significant:

Source of variation	Sum of squares	D.F.	Mean square	
Between years	10076	13	775·08	$F = 2{\cdot}537^{**}$
Within years	265507	869	305·53	
Total	275583	882	—	

** Two asterisks denote significance at the 1% level; a single asterisk indicates significance at the 5% level, and so throughout.

VI. Comparison of pre-war and post-war students

In the attempt to discover the locus of this significant variation, scores were separated into those for pre-war years and post-war years. Analysis of variance was applied to these two groups in turn after tests by Bartlett's method had confirmed that in respect of

dispersion of scores each period was homogeneous. Since the resulting F values (1·51 and 1·42 respectively) are not statistically significant, the pre-war group and the post-war group may each be regarded as having been drawn from a homogeneous population. In other words, the differences between the mean scores over the years within each period could reasonably be ascribed to chance.[1]

The difference of 5·05 between the general means for the two periods (151·26 and 156·31, see Table 2) is, however, highly significant ($t = 3{\cdot}905$; D.F. $= 881$; $P < 0{\cdot}001$) and merits further discussion.

Various explanations can be suggested for this finding:

(*a*) *Age.* The post-war group is in general older than the pre-war group. Before the war the average age of both men and women was about 22 years, while from 1946–7 to 1949–50 the mean age for men was about 26 years, although the women were no older than in pre-war years.

Evidence on the changes of score with age after about 16 on such a test as G.T. 33 is by no means clear-cut. The prevailing opinion would seem to be that with such a highly selected group as ours there will be some increase in mean score well beyond 16. For instance, Dearborn & Rothney (1941) state '. . . our results indicate that mental growth continues up to approximately 30 years, and that approximately 2% of average mental growth may take place after 21 or 22'. With the Moray House Adult Test Thomson (1945) reports 'only slight evidence of an age gradient over the 3 years of non-graduate students'. On the other hand, several investigators (e.g. Miles & Miles, 1932; Wechsler, 1944; Raven, 1948) report a falling off in score after the early twenties—precisely the area with which we are concerned. This is supposed to be especially apparent in speeded tests such as G.T. 33 is. Evidence that we have from supplementary investigations with this test is against the increase in score being due to the increase in mean age of the students. We find, for instance, that there is no significant difference between the mean score (150·82) of another 124 students who did the test at the beginning of their degree course when they were about 19 and that of the pre-war group (151·26) in the present study who were tested after they had graduated (mean age 22 years). Further, if we compare the scores of those men who came to the Department from national service and had a mean age of 26·9 years, with those of men of the usual age for the course, we find that the difference of 5 points in favour of the older graduates falls short of statistical significance. Again, the scores of post-war women are significantly higher than those of pre-war women, although there is little difference in their ages.

(*b*) *Experience.* Though higher mean scores may be found with increasing age in a highly intelligent group, it does not of course follow that the higher scores are due to the greater age *per se*, i.e. that maturation has continued longer at the higher percentile levels. It is very likely that it does. But the greater age means a longer life experience and, with a group of students, longer exposure to a stimulating intellectual environment.

(*c*) *Test sophistication.* The reason for the higher mean age of post-war men students was, of course, the fact that many of them came into the Department after service in the armed forces or in industry. It is probable that many of these had more specific experience likely to affect their scores in the shape of selection or classification tests.

Vernon (1938), and others, have called attention to the effects of familiarity and practice

[1] This is in marked contrast to the findings of Dale (1935), who states 'the data collected during 8 years at Newnham College present an almost startling lack of regularity'.

on scores achieved on intelligence tests. Most researches which have been carried out on experiential effects in intelligence tests have been concerned with comparatively short intervals of time. How long the effect may last has not apparently been investigated. We have ourselves evidence which seems to point to sophistication effects from comparison of test and re-test scores with the same test after an interval of 3 or 4 years. Vernon & Parry (1949) express the opinion that higher scores are likely to be obtained by those who have had experience of tests *even of a different type* as compared with those who are tested for the first time. Is it too far-fetched to suggest that apart from any wartime experience the post-war group are more likely to have come across psychological tests during their schooldays than the pre-war group and that this effect is still operating? It is difficult to believe that this is so but it is not impossible.

(*d*) *Selection.* One plausible explanation of the difference between the pre-war and post-war groups would be that it is caused by a change in standards for admission to the universities, but the way in which this has operated is not at all clear. It would probably be fair to say that in the early post-war period it was relatively easier for ex-service candidates (mostly men) to secure entrance to the university and relatively harder for non ex-service candidates (mostly women), who in many cases had to reach a standard high enough to enable them to enter second-year courses. This, if test-intelligence is related to academic achievement, should be reflected: (i) in higher mean scores for women than men, especially in those schools (mostly in the science faculty) where it was necessary to reach second-year standard; (ii) in a bigger difference between pre-war and post-war women than between pre-war and post-war men; (iii) in a gradual return to pre-war levels of performance as preference to ex-service candidates has become of less influence. The facts are, first, that the difference between men and women in favour of the men is greater since the war than before (though in neither period is the sex difference significant). Secondly, though the mean score for post-war women is higher than for pre-war women, the pre-war–post-war increase of 4·10 (significant at the 2% level) is less than that for the men (6·16). Thirdly, the post-war group has remained homogeneous over the years and there is no sign of a progressive decline to pre-war levels.[1]

All of these factors may have contributed to produce higher scores among the post-war students, and it is not possible to show that any one was mainly responsible, though if we were to hazard a guess it would be that experience of psychological tests, or at least familiarity with 'quizzes' making use of the selective type of answer, was the most likely single influence.

VII. Differences between men and women

The N.I.I.P. norms for grammar-school pupils on this test (N.I.I.P., 1937) give at each percentile level a score for 16-year-old boys 5 points higher than for girls of the same age. It is, however, generally supposed that girls surpass boys at tasks calling for verbal abilities (Terman, 1946). Fitt & Rogers (1950), dealing with a student population very similar to our own, report on Cattell Scale III 'a significant difference in favour of the females between the average male and female I.Q. scores'. Wallace (1952), on the other hand, finds, for universities other than Oxford and Cambridge, that the men make significantly higher scores on Heim's A.H. 4 and A.H. 5 tests though their superiority, she states, is largely due to higher scores on the non-verbal part of the test.

[1] Results from testing 265 students in the 3 years subsequent to the period under discussion confirm that the mean scores are maintained at the higher level.

Results for our men and women students at the different percentile levels can be seen in Table 3. Differences between the means are given in Table 5, and are not significant in either the pre-war or the post-war periods.

Our results on G.T. 33 at the university level do not accord with the findings of Fitt & Rogers on the Cattell Test, nor do they reveal the sex difference reported for G.T. 33 at the grammar-school stage by the N.I.I.P.

The discrepancies may perhaps be due to differential selection. Only the most intelligent of the girls, for instance, may have felt the urge (or been urged) to proceed to a university course.

Many investigations (see Terman, 1946) with unselected groups report a bigger scatter of scores for males. Fitt & Rogers also find a larger standard deviation for I.Q.'s of men

Table 5. *Differences between means and standard deviations of scores of men and women*

	Pre-war		Post-war	
	Men	Women	Men	Women
N	316	305	125	137
M	151·85	150·65	158·02	154·75
Difference	1·20		3·27	
t	0·830 (not significant)		1·636 (not significant)	
S.D.	18·49	17·64	16·41	15·88
Difference	0·85		0·53	
Diff./S.E. diff.	0·848 (not significant)		0·375 (not significant)	

students. They conclude 'it appears that sex differences in central tendency and dispersion can be expected and must be allowed for, when testing gifted adults on the Cattell Scale III'. Table 5 shows that in the present investigation the differences in dispersion of scores are in the expected direction, but are not statistically significant.

Over the years with which this study is concerned, there has been no significant difference at any point of the distribution between our men and women subjects, and if what this test measures is important in teachers, the men and women who have entered the teaching profession from this Department have in this respect been of equally good quality.

VIII. Differences between scores according to class of degree

The numbers of students who obtained honours in special studies, honours in general studies, and pass degrees are given in the following summary:[1]

	Honours, special studies	Honours, general studies	Pass	Fail
Arts	328	15	224	9
Science	145	—	138	14
All	473	15	362	23

In this analysis all students (from other universities as well as Liverpool) were considered since, although standards of different *classes* may well vary from one university to another, yet it is reasonable to assume that there is usually a difference between the main divisions of honours and pass. Honours students in all universities are likely to be similar in ability and more similar to each other than to pass degree students.

[1] These figures do not include 10 students with degrees in Engineering or Commerce.

Taken altogether, the 488 honours students have a mean score of 155·30 against 149·77 for the 362 with pass degrees, and when the results are presented separately for arts and science, as in Table 6, in every case the honours graduates have a higher mean score than those who read for pass degrees (including those who failed to graduate). In the pre-war years the difference is significant at the 1% level for arts, and at the 5% level for science. In the post-war years, also, though they fail to reach statistical significance, the differences are in the same direction.

We now ask if there is a difference in mean score according to class of degree obtained. Here we are concerned with students from Liverpool university only.

Table 6. *Difference between mean score of honours and pass students*

	Arts		Science	
	Honours (incl. gen. hons.)	Pass (incl. fails)	Honours	Pass (incl. fails)
Pre-war				
N	255	157	111	92
M	155·83	148·54	150·47	145·09
Difference	7·29		5·38	
t	4·295**		1·989*	
Post-war				
N	88	76	34	60
M	160·11	156·42	154·68	151·02
Difference	3·69		3·66	
t	1·468 (not significant)		1·072 (not significant)	

Table 7. *Mean score according to class of degree*

	Honours						
	Class I	Class II i	Class II ii	Class III	Gen. hons.	Pass	Fail
Arts, pre-war							
N	48	122	47	1	10	150	6
M	159·83	154·72	153·04	172·00	151·00	149·21	140·00

	Honours						
	Class I	Class II i	Class II ii	Undivided II	Class III	Pass	Fail
Science, all years							
N	27	47	38	9	15	135	14
M	152·44	151·23	151·84	149·56	148·48	147·03	147·28

Arts students. Table 7 shows that there is clearly a trend in the pre-war period for mean scores to drop as we move from first-class honours to fail, and analysis of variance shows that the between-classes variance is significantly higher than that within classes ($F = 4{\cdot}246^{**}$).

On testing the differences in pairs we find within the honours group only one which reaches statistical significance—that between I and II ii (significant at the 5% level). There is a significant difference between first-class honours and other classes of honours combined (significant at the 5% level), but if we take together I and II i (what are usually regarded as 'good honours degrees') against other honours the difference is not significant, though I and II i against all the rest—II ii, III and pass (not including fails)—gives a difference significant at the 1% level.

In the post-war years where the number of students is small no clear pattern emerges.

Science students. Here it is legitimate to combine the pre-war and post-war groups since analysis of variance indicates that they could have been drawn from the same population. Mean scores for the different classes are shown in Table 7. There is a trend of falling score with weaker degree, but the variance of the class means is not significantly greater than the variance within classes.

If as a matter of interest we lump together the scores of all arts and all science students from Liverpool University and group them according to class of degree obtained the results are as follows:

	Class					
	I	II i	II ii	III	Pass	Fail
Number	81	213	108	19	352	23
Mean score	157·05	155·30	153·70	151·37	149·84	147·83

The general trend for these 796 Liverpool graduates who have passed through the Department of Education is sufficiently clear: the means decrease as we pass from first class honours degrees to failures.

Using the same test, Babington Smith (1940) reports a mean score of 154·7 for first- and second-class honours combined and a mean of 149·3 for third-class honours and ordinary degrees combined. For our pre-war students the corresponding figures are 155·22 and 149·92. Dale (1935) gives 157·3, 155·1 and 151·4 as mean scores for first-class, second-class and third-class honours in the Cambridge Tripos examinations for her Newnham College subjects.

IX. Differences between scores of arts and science students

Since G.T. 33 is so obviously a 'verbal' test, it might reasonably be expected that those who have had a predominantly literary training at the university would have an advantage over those who have pursued a scientific course. Our results certainly show that both in pre-war and post-war years the mean score for arts graduates is greater than that for science graduates to a highly significant degree.

	Pre-war				Post-war			
	N	*M*	Diff.	*t*	*N*	*M*	Diff.	*t*
Arts	417	152·90	4·99	3·256**	167	158·51	6·07	2·956**
Science	204	147·91			95	152·44		

Whether these results mean that the test is relatively easier for arts students or whether selective factors are again responsible cannot be decided.

If we apply the same kind of analysis that we used for all students in section V to arts students alone we find that there is a highly significant difference between the means over the years ($F = 2{\cdot}622^{**}$), though the pre-war and post-war groups are themselves homogeneous. The difference of 5·61 between the two groups is highly significant. For men arts students the difference reaches the highly significant value of 8·01, while for women the difference of 3·84 is in the same direction though it is not significant.

Science students form a group which is homogeneous over the years, but the mean score of post-war students is 4·53 higher than for pre-war, the increase being similar for both men and women.

The superiority of arts over science graduates was analysed further by considering men, women, honours, and pass students separately. In each case the scientists obtained

a lower mean score than the arts students in the corresponding group—in five out of the eight comparisons the differences are significant, two of them at the 1% level.

These results with G.T. 33 may be contrasted with the results we have obtained using the N.I.I.P. non-verbal group test 70/23 on which science students tend to do better than the arts. Other investigators, e.g. Vernon (1939), Heim (1947), Foulds & Raven (1950), Wallace (1952), have reported that students with a science background tend to gain higher scores on non-verbal tests while the arts students do relatively better on the verbal type. Vernon, for instance, states that the results of the verbal and non-verbal tests are 'decidedly different' for arts and science students. He also notes that university graduates scored more highly than non-graduates on G.T. 33, while there was no such difference on the Stephenson non-verbal test which he used. 'This suggests', he says, 'that the two groups may be of equal *g* but that verbal tests are markedly affected by university training.' It would appear probable from our results that they are more affected by an arts than by a scientific training, and that in any attempt to compare highly intelligent and highly educated people the medium through which 'intelligence' is tested should be taken into account.

X. Differences between scores according to subjects of degree

Here we group the students according to their main field of study, confining ourselves to those who obtained honours in special studies, since it would be too complex a task to attempt to classify those with general honours and pass degrees. Only Liverpool graduates are considered. These restrictions mean that the numbers in some of the groups are very small, and in some cases we have combined two or more groups. The classification will be clear from Table 8, which shows the mean score from each subject group for the pre-war period and for the pre-war and post-war periods combined.

Table 8. *Mean scores of honours students according to subject of degree*

	Pre-war		Pre-war + Post-war	
	N	*M*	*N*	*M*
Arts subjects				
Classics	26	160·04	36	162·94
Modern Languages (French, German, Spanish)	59	159·64	74	159·49
History	54	154·96	67	155·39
English	37	151·89	59	156·48
Geography	42	150·36	57	149·98
Science subjects				
Mathematics	32	158·38	35	157·71
Biological Sciences (Botany, Zoology)	36	147·39	43	148·84
Physical Sciences (Physics, Chemistry)	37	146·78	57	148·67

Analysis of variance over these groupings gives a significant value of F, and the results of tests for the significance of differences (t values) between mean scores for subjects are shown in Table 9. Among the arts students it will be seen that Classics students do significantly better than English and History students at the 5% level. The scores for both Classics and Modern Languages are higher than the mean for Geography to a highly significant degree. (These conclusions, based upon results for pre-war and post-war periods combined, are essentially the same as those arrived at from a consideration of pre-war figures alone.)

With the science students, a test of the differences in pairs confirms the impression that the mean score for Mathematics is significantly higher than those for the physical sciences and for Biology. The difference between these latter groups is not significant.

Inspection of the whole pattern without regard to faculty shows that of honours graduates from Liverpool University, the Classics, Modern Languages and Mathematics groups have the highest scores in that order, Geography, Biological Sciences and Physical Sciences the lowest, with English and History holding intermediate positions.[1] Wallace (1952) on Part 1 of A.H. 5 also found highest mean scores for Mathematics, Classics and Modern Languages, and the ranking of other subjects is in the main the same as in our table. Dale (1935) reports that of her Newnham students highest marks on a combination of G.T. 33 and the Cattell Test were gained by Mathematics and Classics followed by Natural Science, English, Modern Languages and History. There is, therefore, a considerable measure of agreement between our present results and those of similar investigations.

Table 9. *Significance of differences between mean scores for subjects* (t *values*)

Subject	*N*	*M*	Modern languages	Mathematics	English	History	Geography	Biological Sciences	Physical Sciences
Classics	36	162·94	1·124	1·486	2·020*	2·474*	3·845**	3·813**	3·765**
Modern Languages	74	159·49	—	0·550	1·095	1·571	3·321**	3·337**	3·483**
Mathematics	35	157·71	—	—	0·370	0·725	2·186*	2·277*	2·290*
English	59	156·48	—	—	—	0·393	2·136*	2·262*	2·347*
History	67	155·39	—	—	—	—	1·873	2·041*	2·131*
Geography	57	149·98	—	—	—	—	—	0·324	0·381
Biological Sciences	43	148·84	—	—	—	—	—	—	0·044
Physical Sciences	57	148·67	—	—	—	—	—	—	—

Classics and Mathematics, which head our list, were for long regarded as the most valuable subjects for the purpose of intellectual training. If psychology can no longer lend support to a belief in their special efficacy, it is nevertheless probable that many head teachers believe in their supreme educational value, and there is little doubt that many schools tend to 'put in' their ablest pupils for one or the other course. Equally, there are unquestionably some subjects, e.g. Geography and Biology, which are regarded (with or without justification) as 'soft options' in which a safe pass in the first school examination may be expected for the mediocre pupils. Thus it may really be selective factors which are at work and junior courses may act as sieves to sort out the more intelligent.

At the university level, selection may operate further to differentiate between the different subject groups. It is worth noting in our results that though geographers are the lowest scorers on the arts side there is no evidence that biologists are at all inferior to the Physics and Chemistry students. One possible explanation of this may lie in differences of employment prospects after graduation in that the more able physicists and chemists may have entered better paid posts in industry and not have come to the Department of Education to train for the teaching profession. In addition, quite apart from any theories of formal discipline, the nature of the course followed may affect the marks gained on an intelligence test. It is possible, for instance, that a linguistic background or mathematical training may give their possessors advantages in solving the sort of problems commonly used in intelligence tests. But the reasons for these subject differences are probably multiple and complex, and no single explanation is likely to be adequate.

[1] Geography, which at Liverpool is only taken in the arts faculty, is, of course, in many ways more akin to the science courses than the other arts subjects considered.

XI. Summary

Results on the N.I.I.P. Group Test 33 for some 900 students in a University Department of Education over a period of 18 years are presented. Analysis shows consistent levels over the pre-war and the post-war years, and some interesting differences between the two periods. Students in the post-war group make significantly higher scores than those pre-war. There are some significant differences between the performances of arts and science graduates and between those with honours and pass degrees. The difference between men and women is not significant. Differences according to class of degree and main field of study are examined. These results are discussed and compared with the findings of other investigations with university students. A table of percentile norms is given.

References

Dale, A. Barbara (1935). The use of mental tests with university women students. *Brit. J. Educ. Psychol.* v, 59–75.

Dearborn, W. F. & Rothney, J. W. M. (1941). *Predicting the Child's development.* Cambridge, Mass.: Sci-Art.

Fitt, A. B. & Rogers, C. A. (1950). The sex factor in the Cattell Intelligence Tests, Scale III. *Brit. J. Psychol.* xli, 186–92.

Foulds, G. A. & Raven, J. C. (1950). An experimental survey with Progressive Matrices. *Brit. J. Educ. Psychol.* xx, 104–10.

Heim, A. W. (1947). An attempt to test high-grade intelligence. *Brit. J. Psychol.* xxxvii, 70–81.

Miles, C. C. & Miles, W. R. (1932). The correlation of intelligence scores and chronological age from early to late maturity. *Amer. J. Psychol.* xliv, 44–78.

National Institute of Industrial Psychology (1937). Manual of Instructions for Group Test 33.

Raven, J. C. (1948). The comparative assessment of intellectual ability. *Brit. J. Psychol.* xxxix, 12–19.

Slater, P. (1940). Some group tests of spatial judgment or practical ability. *Occup. Psychol.* xiv, 40–55.

Smith, B. Babington (1940). The conversion of scores on Group Test 33 to intelligence quotients. *Occup. Psychol.* xiv, 184–97.

Snedecor, G. W. (1940). *Statistical methods.* Ames, Iowa State College Press.

Terman, L. M. (1946). Psychological sex differences. In Carmichael, L.: *Manual of Child Psychology.* New York: Wiley.

Thomson, G. H. (1945). The distribution of intelligence among university and college students. *Brit. J. Educ. Psychol.* xv, 76–9.

Vernon, P. E. (1938). Intelligence test sophistication. *Brit. J. Educ. Psychol.* viii, 237–44.

Vernon, P. E. (1939). Educational abilities of training college students. *Brit. J. Educ. Psychol.* ix, 233–50.

Vernon, P. E. & Parry, J. B. (1949). *Personnel Selection in the British Forces.* London: University of London Press.

Wallace, Jean G. (1952). Results of a test of high grade intelligence applied to a university population. *Brit. J. Psychol.* xliii, 61–9.

Wechsler, D. (1944). *The Measurement of Adult Intelligence.* Baltimore: Williams and Wilkins.

(*Manuscript received* 14 *January* 1953)

INFLUENCES AFFECTING THE RESULTS OF A TEST OF HIGH-GRADE INTELLIGENCE

By KATHLEEN P. WATTS

Psychological Laboratory, University of Cambridge

I. *Introduction* (p. 359). II. *Effects of practice* (pp. 359–360). III. *Speed and power* (pp. 360–361). IV. *Careless errors* (pp. 361–362). V. *Differences in score between the sexes* (pp. 362–363). VI. *Differences between the faculties* (pp. 363–364). VII. *Examination classes* (pp. 365–366). VIII. *Summary and conclusions* (pp. 366–367). *References* (p. 367).

I. Introduction

In the course of performing an item analysis of a high-grade intelligence test (AH 5), it was found that some points of general interest emerged, in addition to data specifically about the test itself. The results of the item analysis indicated that the individual test questions, and their order, were satisfactory: this paper is concerned only with the more general points, such as the (short-term) effects of practice, the interaction of 'speed' and 'power', the relation between intelligence test scores and careless errors, and the intellectual differences between men and women which may be deduced from a comparison of their test scores and of their examination results.

Test AH 5 is a high-grade intelligence test in two parts: part I has twenty-three verbal and thirteen numerical questions, and part II thirty-six diagrammatic questions. The time limit of 20 min. set for each part of the test is, by design, too short for most subjects to attempt every question. The test was given to 767 (volunteer) students from eight universities or colleges. Some of these had taken test AH 4 shortly before taking test AH 5.[1] AH 4 is a relatively easy and short test; but it is similar in form to AH 5 and has the same biases. For this reason, it was thought that practice in AH 4 might have some effect on performance in AH 5. The results were, therefore, analysed in two groups: group A who took AH 5 only (484 students) and group B who took both AH 4 and AH 5 (283 students).

II. Effects of practice

We found a difference between groups A and B when we considered the number of questions reached (the latest question in each part for which an answer had been entered was taken as the last question *reached*). The group who had taken AH 5 only (group A) reached 81% of the questions in part I and 86% of the questions in part II. These percentages were higher for the group who had taken AH 4 before taking AH 5 (group B), being 88% in both parts. In group B, only one subject did not reach Question 20 in part I, although thirty of group A failed to reach this question. The discrepancy between the two groups is smaller in part II—seven of group A and four of group B failing to reach Question 20.

The group who had recently taken the easier test (AH4) may have inferred that it is desirable to work quickly on any timed intelligence test, but the other group may not have thought about speed of working until they observed that they failed to complete

[1] Tests AH 4 and AH 5 will shortly be available to accredited users on application to the Psychological Laboratory, Cambridge.

part I of AH 5 in the given time. So we find that subjects in group A reach fewer questions in part I, but in part II the number they reach approaches the number reached by group B in both parts of AH 5.

This is shown also by the percentage of questions *answered*—group A increasing from 66% answered in part I to 74% in part II, and group B increasing only from 72 to 76%. Thus group A improved more in the number answered than did group B, thereby suggesting that they genuinely increased their speed and were not reaching more questions by simply omitting the harder ones.

III. Speed and power

It has been demonstrated that, when applied to university students, Test AH 4—an intelligence test designed for a cross-section of the entire population—tends to differentiate on speed rather than intelligence, and the Test AH 5, which is designed for a population of high-grade intelligence, does differentiate on intelligence—or, at least, on something other than speed (Heim, 1947; Heim & Batts, 1948). These results were obtained when both tests were given first with limited and then with unlimited time. To see if the present data could yield any further information about the role of 'speed' and 'power' in these tests, we examined the results of the 283 students (group B) who had taken both tests in limited time.

We could not determine the actual 'speed' at which a subject worked, i.e. the number of questions which he *attempted* in the time set; but we knew it could not be less than the number answered nor greater than the number reached. We therefore calculated, for both parts of the test, the mean score of those who (i) answered n questions, and (ii) reached n questions.

With AH 4, the subjects are instructed to work through the questions in order, and so the number reached is almost identical with the number answered. On the average, subjects answered fourteen-fifteenths of the questions correctly in part I and only slightly less in part II. This high proportion of correct answers is maintained by subjects of all speed-levels. It would seem, therefore, that the gradient of difficulty is so gradual, and the later questions so easy, that speed makes little difference to the accuracy of the answers given by university students. There is little difference between the men's and women's results on this test.

In test AH 5 roughly three out of four questions *answered* are correct (a rather bigger proportion for the men and a rather smaller one for the women). Many subjects who approach the end of the test do so only by omitting questions; thus the ratio of *correct to reached* gradually becomes smaller than that of *correct to answered*. It was observable that this trend begins about Question 28 for the men and for the women about Question 24—as they tend to omit more questions than do the men.

Thus it would seem that the quick and intelligent subjects approach the end of the harder test and gain relatively high scores; but the quick subjects who are not so intelligent find the general level of difficulty so high and the gradient so steep that, although they manage to approach the end of the test, they gain no more marks than the subjects who work slowly and steadily through the earlier questions.

The scatter of scores about these means is greater for AH 5 than for AH 4—particularly in part I of the tests. For instance, the twelve men who reached the end of part I of AH 5 had scores ranging from 9 to 34 (maximum score 36) on that part of the test, whereas the

sixteen men who reached the end of part I of AH 4 had scores ranging from 55 to 65 (maximum score 65). These results suggest that although speed of work plays some part in determining scores on both tests, when they are applied to highly selected intelligent subjects, AH 5 scores are influenced far more than the AH 4 scores by other factors.

IV. CARELESS ERRORS

Intelligence tests are sometimes criticized (though evidence is rarely adduced) on the grounds that the quick facile workers tend to gain the highest scores, although they are liable to be more careless in general than subjects who work more slowly. It was hoped that this analysis would throw some light on this problem.

We picked out the nine questions where, in our opinion, it was possible to make a careless error—for instance, where the subject gave the (correct) *next member* of a series when the *next member but one* was requested; where he omitted the decimal point, having calculated the digits correctly; or where he put given words in (the right) order but started with the smallest instead of with the largest as requested. We found that one or more of these careless errors had been made by over 70% of the subjects. The number of careless errors made by individuals varied from 0 to 6, with a mean of 1·25 for group A and 1·42 for group B.

Table 1. *Number of careless errors and mean AH 5 total scores*

Number of careless errors	Group A		Group B	
	No.	Mean	No.	Mean
0	146	38·36	64	40·19
1	170	38·56	100	39·53
2	98	39·83	68	40·03
3	46	40·57	37	40·89
4	19	38·74	10	39·50
5+	5	39·80	4	42·50

As six of the nine questions were numerical, it was thought that our findings might apply to arithmetical carelessness only. However, when these nine questions were examined individually, it was found that in each case the subjects who give a careless answer to these questions have a higher mean total score than those giving any other wrong answer or omitting the question.

The numbers of careless errors subjects made were compared with their rates of work (as measured by the number of questions reached). Though there seemed to be a tendency for the quicker subjects to make more careless errors than the slower subjects, the difference between the distributions of number of careless errors at different levels was not significant (group A, $\chi^2 = 8{\cdot}83$ with 8 D.F.; group B, $\chi^2 = 6{\cdot}50$ with 6 D.F.).

If the only difference between 'careless' subjects and 'careful' subjects lay in their degree of carelessness, subjects making, say, five careless errors might be expected to score about four marks less than those making one careless error. It can be seen from Table 1 that this is not the case. The scores are all roughly equal. (The figures show, if anything, a slight upward trend, but this may be due to the very slow workers (with inevitably low scores) attempting fewer questions.)

It seems probable, then, that if quicker subjects are more careless than slower subjects, their speed and grasp compensate (perhaps more than compensate) for loss of marks by carelessness. For instance, group B, which works faster than group A (see Effects of

Practice), makes slightly more careless errors than group A; yet, even so, group B has a higher mean total score.

To make a careless error, as defined above, it is necessary to grasp the principle of the question, and the duller the subject, the more likely he is to give a wrong answer, or to omit the question, through failure to comprehend what is required. It is to be expected, therefore, that the duller subjects will make less careless errors than will the more intelligent subjects; inability to comprehend, or a very slow rate of work, will affect the number of careless error questions available to the subject. We should suggest that previous generalizations made concerning lack of carelessness on the part of the slower and duller subjects, may have neglected the fact that the commission of an 'uncomprehending' error on a question is usually incompatible with the commission of a careless error on that question.

V. Differences in score between the sexes

It can be seen from Table 2 that the mean scores for the men are higher than those for the women in both groups (significant at the 1 % level for group A and 0·1 % level for group B). In Table 3, which appears in the next section, it may be seen that AH 5 score varies with

Table 2. *Mean scores for men and women*

		Men ($N=373$)	Women ($N=111$)
Group A	Part I	18·95	17·97
	Part II	20·57	19·13
	Total	39·52 (S.D. 8·44)	37·10 (S.D. 7·46)
		Men ($N=120$)	Women ($N=163$)
Group B	Part I	20·88	19·10
	Part II	21·72	19·02
	Total	42·60 (S.D. 8·41)	38·10 (S.D. 7·52)

Differences between the mean scores of the two groups (B – A)

	Men	Women
Part I	1·93	1·13
Part II	1·14	−0·11
Total	3·08	1·02

Faculty. As the proportion of students from each Faculty is different for men and for women, this may account for some of the difference between the mean scores for the two sexes. However, the men score more highly than the women in *every* Faculty (except the two male Social Studies students in group B, whose very low part II scores suggest that they were not representative of their Faculty). It would appear then, that on the whole, the women found test AH 5 more difficult than did the men.

We have shown earlier in this report that practice has produced an increase in *rate of work* between parts I and II of the test with group A, and on both parts of the test with group B. As can be seen in Table 2, the same pattern applies to the *scores* of both sexes. The men evidently gain more benefit from practice on a similar test, since the difference between the mean total scores of the two groups is significant at the 0·1 % level for the men though, on these numbers, it is not significant for the women.

It has been suggested that women do relatively better in part I, whereas men do relatively better in part II. It can be seen from Table 2, that whilst there is a general superiority of men over women, the greater difference is between the part II scores. The groups were split into male and female subgroups of ten (or over) on the basis of total

score on the test, and the average differences between parts I and II for these subgroups were calculated. The range of scores for men and women were roughly the same, rendering possible a straightforward comparison between the subgroups. For group A, in every male subgroup the average part II score was greater than the average part I score, but this was not true for the female subgroups. Here the high-scoring subgroups on the average gained more marks on the diagrammatic part of the test, but the low-scoring subgroups gained more marks on part I. Thus it would seem that the more intelligent women, like the men of all levels, do relatively better on the diagrammatic part, and that the tendency for women to do better on part I is confined to the *less intelligent* women. A similar result was found in analysis of group B, where they followed the same pattern, although the differences were smaller.

On the whole, subjects gain slightly higher marks in part II than in part I of the test (with the exception of the low-scoring women, as mentioned above), suggesting that the diagrammatic part is relatively easier than the verbal and numerical part. This relative ease may be due to the biases of the parts, or it may be due to the fact that the diagrammatic part is the *second* part of the test. We intend to test a group of students on AH 5 giving part II first to see whether the difference in score between the two parts is attributable mainly to practice.

In both groups the women tended to get further through the test than the men, but they omitted more questions in so doing. Men and women of all intelligence levels found the verbal questions easier than the numerical or diagrammatic questions, and the women scored as highly as the men on the verbal questions only. Numerical questions were the hardest type to answer, and only men and women in the top decile solved more than half of them correctly—though, in general, more were omitted than answered incorrectly.

VI. Differences between the faculties

As in an earlier report on these data (Wallace, 1952), the groups were divided into four main 'Faculties':

(*a*) *Arts:* including English, Modern Languages, Classics, History, Law and degrees such as the Scottish Ordinary M.A. which include a varied combination of subjects.

(*b*) *Mathematics:* including Engineering, Natural Philosophy, Astronomy and Mechanical Sciences.

(*c*) *Science:* including Medicine, Physics, Chemistry, Agriculture, Pharmacy, Physiology, Zoology, Botany and Biochemistry.

(*d*) *Social Studies:* including Economics, Geography, Moral Sciences, Psychology, Commerce and Anthropology.

This is only a rough classification, but it is useful for purposes of comparison as it gives us four general groups into which all but eight of the students can be satisfactorily fitted. The mean scores for students of these four Faculties are shown below in Table 3.

The Arts students—with the exception of women in group A—gained more marks in part I than in part II of the test. The other students gained more marks in part II (except for the two male students of Social Studies in group B). As might be expected, Mathematics students scored more highly in the numerical and diagrammatic questions and the Arts students more highly in the verbal questions.

If, for each Faculty, we consider the difference between the mean part I score of group A

and the corresponding part I score for group B we find that only with the Arts Faculty is the difference greater than 2·0 marks. (The differences between the part II scores are also less than two marks for all cases except the female students of Social Studies.) This suggests that the Arts students benefit most from practice in a similar test. If, as we have suggested, the order in which the parts of the test are taken makes a difference with subjects who have little or no previous test experience, the test in its present form may penalize the Arts students.

When the ranges of score of the Faculties were compared it was found that the upper limits were roughly the same. At the lower end, however, none of the Mathematics students gained less than 30 marks on total score, although several students from each of the other Faculties had scores well below this mark. As the number of students of the Mathematics Faculty is smaller than that of the other three Faculties, this may be simply due to sampling. On the other hand, it may be that the minimum amount of 'intelligence' required to study Mathematics (and allied subjects) is higher than the minimum required to study in other Faculties; the types of problems which occur in AH 5 may have a greater appeal to Mathematics students; there may be transfer of training between

Table 3. *Faculties and mean scores*

	Group A. Men				Group B. Men			
Faculty	No.	Pt. I	Pt. II	Total	No.	Pt. I	Pt. II	Total
Arts	76	19·70	19·17	38·87	60	22·03	21·15	43·18
Mathematics	40	19·93	23·07	43·00	10	20·00	22·90	42·90
Science	187	18·53	20·90	39·43	48	19·67	22·40	42·06
Social studies	68	18·88	20·09	38·97	2	20·00	16·50	36·50
	Group A. Women				Group B. Women			
Arts	38	17·61	18·32	35·92	83	19·67	17·78	37·46
Mathematics	—	—	—	—	7	19·43	23·43	42·86
Science	45	18·22	19·73	37·96	42	18·31	19·33	37·64
Social studies	27	17·74	19·07	36·81	26	18·50	21·35	39·85

Mathematics and intelligence test questions; or perhaps a combination of all these factors influences the test scores of the Mathematics students.

Wallace (1952) found that 'those students studying Mathematics and Engineering scored significantly higher marks on AH 5 than students in any other Faculty groups, except on part I, where there was no significant difference between this group and those studying Arts subjects'. This finding was based on much the same data as the present analysis and so it would seem that the difference found with the Mathematics group was due partly to the proportion of men subjects to women subjects. If we compare *all* the Mathematics students included in this analysis with *all* the Arts students we find that the Mathematics students have a mean total score of 42·96 and the Arts students a mean score of 39·37. This difference is significant at the 5% level. However, as the ratio of men to women is 50 : 7 for the Mathematics Faculty compared with 136 : 121 for the Arts Faculty, much of this difference may be due to *sex* differences and not to differences between the Faculties.

When we take into consideration the effects of practice, the effects of bias, the differences between the mean scores of men and women (and there may also be a university effect), it is not possible to draw any definite conclusions as to whether members of one Faculty are 'more intelligent' than those of another.

VII. EXAMINATION CLASSES

We thought it would be interesting to see if there were any differences between the average *Examination Classes* gained by students of different faculties and, if so, how they compared with the results of this analysis.

We took the Cambridge Tripos Examination (parts I and II) over five evenly-spaced 5-year periods between 1920 and 1952 and calculated the percentage of students gaining, respectively, First, Second and Third Classes or (failing to reach this standard) 'Specials', in Mathematics, English and Natural Sciences. On the whole, the proportions in each Class for a particular Examination were fairly constant over the five periods. For this reason, only the mean percentages for the 25 years (as a whole) are shown in Fig. 1.

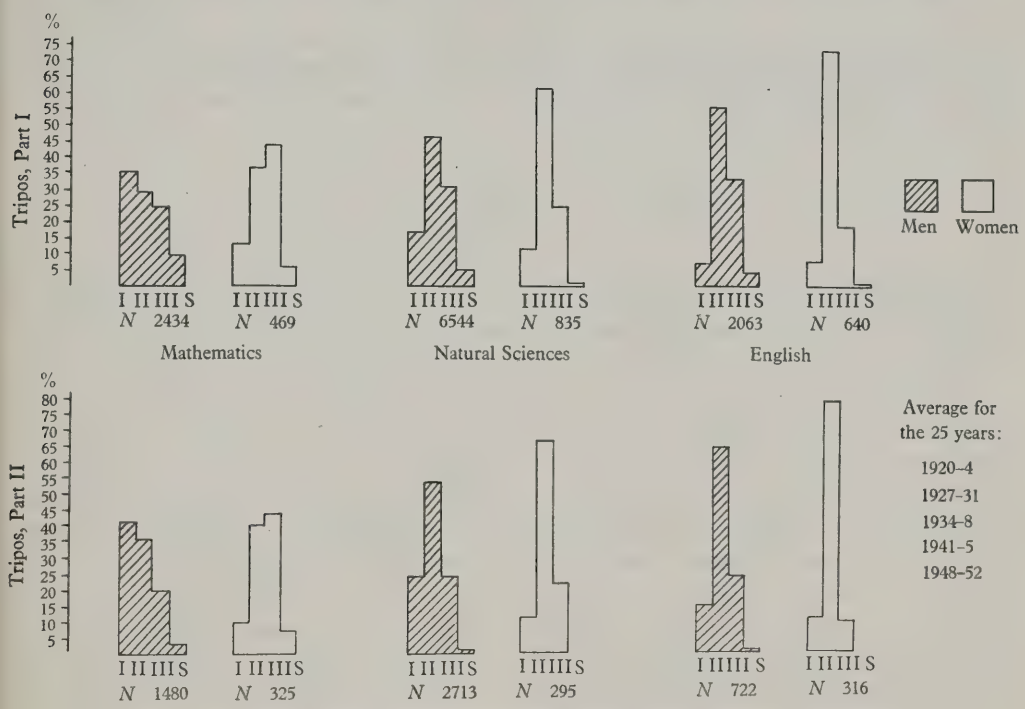


Fig. 1. Comparison of the percentage of First, Second, Third Classes and 'Specials' awarded in the Cambridge Tripos Examination.

It is difficult to say how much the standard of marking varies from one faculty to another. For instance, Mathematics papers can be marked more objectively than English papers: examiners are more likely to be in agreement as to the correctness of a mathematical equation than they are on the merits of a certain style in an English essay. But the standard of marking is identical for the two sexes, and so the differences found between the Examination Classes gained by men and by women may be presumed to be due to psychological differences. (As the admission of women undergraduates at Cambridge is more highly competitive than that of the men they might have been expected to gain, if anything, better Classes than the men.)

The most striking difference between the Classes gained by the male and by the female students is in Mathematics, Firsts being the Class most frequently gained by the men and

Thirds being the Class most frequently gained by the women. This supports the idea that men are naturally 'more mathematically-minded' than women, although differential mathematical training during the years at school may well play some role. With AH 5 we found that the men were better than the women in both the numerical and diagrammatic questions.

Women gain as many Firsts in Mathematics as they do in Natural Sciences or English; but they obtain a far higher proportion of Thirds and Specials in Mathematics than in the other two faculties. On AH 5 the high-scoring women found the diagrammatic part of the test no harder than the numerical and verbal part; but the low-scoring women found the diagrammatic part relatively difficult. It would seem then that Mathematics and diagrammatic questions are relatively difficult only for the less intelligent women.

In the English and Natural Sciences Triposes, the men gain more Firsts *and* more Thirds than the women, the women gaining a disproportionate number of Seconds. Thus, in this as in other fields, the women tend less towards the extremes than the men. This was found also with the intelligence test scores of the two groups used in this item analysis, where the standard deviations were as follows:

	Men	Women
Group A	8·44	7·46
Group B	8·41	7·52

Women do better than men in part I of the English Tripos, and at least as well in Part II, English Tripos. Women gain fewer Thirds in English (Parts I or II) than in Natural Sciences or Mathematics, whereas men gain the most Thirds in English Part I. Women were found to score as highly as men on verbal questions in AH 5.

The difference between the results of Part I and Part II of the Tripos Examination is much more marked for the men than for the women. In all three faculties, the men gain a larger percentage of Firsts and a smaller percentage of Thirds and Specials in Part II of the Tripos than they do in Part I. The women, in Part II, gain more Firsts in English, but they gain *fewer* Firsts in Natural Sciences and Mathematics. They get the same percentage of Thirds and Specials (taken together) in both parts of the Mathematics Tripos, but show some decrease—though less than the men—in the percentage of Third classes gained in Part II of the English and Natural Sciences Tripos. It is interesting (though the two cases are not, of course, analogous) to notice that in AH 5 the men gained more benefit than the women from short-term practice.

It would seem, then, that some psychological differences between men and women may be deduced from a study of examination results and from a study of intelligence test scores; and that the two situations evoke similar trends.

VIII. Summary and conclusions

1. A high-grade intelligence test, AH 5, was given to two groups of (volunteer) students from eight universities and colleges. Group A (484 students) took AH 5 only and group B (283 students) took AH 5 shortly after taking an easier intelligence test, AH 4. The usual time limits were imposed.

2. Practice (in AH 4 for group B and part I of AH 5 for group A) produced an improvement in test performance, which was due partly to an increase in speed of work.

3. The results of the easier test (AH 4) were compared with those of the more difficult test (AH 5) for group B. With AH 4, the quickest subjects gained the highest scores, but

this was not necessarily true with AH 5. Although the role of speed cannot be ignored, other factors play a large part in determining AH 5 scores.

4. Errors were defined as 'careless' when it was clear that the subject had grasped the logical principle underlying a test question but had failed to observe the exact instructions. It was found that careless errors were made by most subjects, and that if the quicker subjects made more careless errors, by their speed and grasp they compensated for marks lost through carelessness. The duller and slower subjects contributed more than their quota of 'uncomprehending' errors and omissions and, therefore, made proportionately fewer 'careless' errors.

5. The men tend to gain higher scores than the women on AH 5. They also gain more extra marks with short-term practice. Only the less intelligent women (judged by score on the test) found the diagrammatic part of the test more difficult than the verbal and numerical part.

6. For purposes of comparison, the students were grouped into four main 'Faculties'—Arts, Mathematics, Science and Social Studies. The lower scorers in Mathematics gained considerably more marks than the lower scorers in the other three Faculties. It is suggested that most apparent differences between the test scores for the Faculty groups may be due to the effects of practice, the effects of bias and the differing proportions of male and female students rather than to differences in the average level of intelligence of the members of the different Faculties as such.

7. It was found that a study of the Classes gained by men and women respectively in the Cambridge Tripos Examinations revealed some intellectual differences between the sexes, and that these differences are similar to those found on high-grade intelligence tests.

Test AH 5 was devised by Dr A. W. Heim and this inquiry was planned at her suggestion. The writer would like to thank Dr Heim for her help in writing this report, and Miss V. R. Cane for some help with the statistics. She would like, too, to express her thanks to the Medical Research Council for their financial support.

REFERENCES

HEIM, A. W. (1947). An attempt to test high grade intelligence. *Brit. J. Psychol.* XXXVII, 70–81.

HEIM, A. W. & BATTS, V. (1948). Upward and downward selection in intelligence testing. *Brit. J. Psychol.* XXXIX, 22–9.

WALLACE, JEAN G. (1952). Results of a test of high-grade intelligence applied to a university population. *Brit. J. Psychol.* XLIII, 61–9.

(*Manuscript received* 17 *February* 1953)

THE STRUCTURE OF VERBAL FLUENCY[1]

By CYRIL A. ROGERS[2]

Institute of Education, University of London

I. *Introduction* (p. 368). II. *Rationale* (pp. 368–372): (i) *The problem* (p. 369); (ii) *Methods of measuring fluency* (p. 369); (iii) *The population and sample* (pp. 369–372); (iv) *The battery of measures* (pp. 370–372); (v) *Reliability and validity* (p. 372). III. *A factor analysis of the data* (pp. 372–378): (i) *The extraction of common factors* (pp. 372–374); (ii) *The rotational problem* (pp. 374–375); (iii) *Interpretation of the factors* (pp. 375–378); (iv) *Hypotheses and conclusions* (pp. 378–379). *References* (pp. 379–380).

I. Introduction

Toward the end of last century Sir Francis Galton (1883) introduced the term 'fluency' into psychological literature. Although he does not state so explicitly, it is evident that by fluency he means the facility with which ideas can be called from the 'antechamber of consciousness'—roughly equivalent to the subconscious in later terminology—into full consciousness. Two forms are briefly mentioned, namely, literary fluency and visual fluency, and Galton points to a possible relation between them.

Spearman's (1904) introduction of the two-factor theory of intelligence, however, eliminated by definition the special fluency factors which had been suggested by Galton, and this position was supported by Carey (1916) and Davey (1926), two of Spearman's students. From a different factorial point of view, Woodrow (1939) and Cureton (1947) also reject the hypothesis of a verbal fluency factor.

On the other hand, Burt's (1917) pioneer report to the L.C.C. clearly uncovers two verbal factors within the cognitive area of personality, one of them displaying a definite fluency component. Similar findings have been reported by Hargreaves (1927), Cattell (1936), Thurstone (1938), Holzinger & Harman (1938), Eysenck (1939) and Bechtholdt (1947). Other studies suggest that verbal fluency may be more complex than a single factor; for example, Carroll (1941 *a*) splits it into two or possibly three factors and so does Taylor's (1947) study.

Because of the diversity of the findings, we intend to sample the different types of tests and factors, and analyse them through a common matrix of intercorrelations. In this way it should be possible to arrive at a more secure description of the structure of verbal fluency.

II. Rationale

This investigation draws on a number of statistical procedures, the main ones dealing with factor analysis. The three chief functions of statistics have been outlined by Kelley (1947): it is descriptive, it enables one to invent a hypothesis and it assists in the verification or disproof of a hypothesis. The main portion of this cognitive exploration of fluency is concerned with description and the construction of hypotheses. It would appear that the most parsimonious interpretation of factors, with a maximum explanation of the original

[1] Based on a paper read to the British Psychological Society—Education section—on 24 May 1952.

[2] I should like to express my appreciation to Prof. P. E. Vernon, for his generous encouragement and helpful criticism. The investigation was made possible by an Imperial Relations Trust Fellowship, and a grant from the Central Research Funds Committee of the University of London.

correlational matrix, can be achieved within a hierarchical framework. Consequently this position, as exemplified by Vernon (1950) and Eysenck (1947), has been adhered to.

(i) *The problem*

A selection has been taken of those fluency tests used in the normal, neurotic and psychotic fields, so that the factors of any significant variance are represented by tests in our battery. Particular attention has been paid to the nature of oral fluency, and tests of it have been included to parallel factors discovered in the written fluency field.

Finally, an attempt is made to set up hypotheses in answer to the following questions:

(*a*) Does the major variance of a battery of fluency tests fall on a factor or factors, distinct from the usual *v* or *V* of British and American researches?

(*b*) How many fluency factors can be identified with confidence and what is their essential nature?

(*c*) What is the most parsimonious explanation of fluency?

(*d*) How much emphasis can be given to the numerous factors established by American researches when they are investigated from a hierarchical point of view?

(ii) *Methods of measuring fluency*

The task of measuring written fluency presents no difficulty as the usual pencil-and-paper methods are applicable. On the other hand, the task of accurately recording oral fluency is one of considerable complexity. The methods of recording it boil down to three: the verbal material can be taken down in shorthand or longhand by the examiner; a third person can be employed inside the testing room, or behind a one-way screen; some form of sound recording can be used.

The first two methods have been used in other investigations, but both have serious disadvantages. For example, it is difficult to make proper observations of a childs' behaviour and reactions when continuously writing; it often distracts the child, and there is no definite method of checking on the accuracy of the material taken down. Furthermore, with rapid or slurred speech, even shorthand becomes inadequate at times. The presence of a third person in the testing room is obviously damaging to the establishment of rapport, and the technique of using a one-way screen is out of the question when travelling from school to school. After weighing the various problems involved, we decided to use sound-recording for the oral tests.

As a result of several trial runs with various kinds of wire, tape and cylinder recorders, we selected a Dictaphone as being the most suitable for our purposes. The actual recordings were made on small plastic disks, each retaining a total of 15 min. These recordings were easily storable and could be played-back any number of times. To facilitate recording we constructed a special desk microphone, omnidirectional up to a range of 5 ft. This allowed satisfactory recordings at close range, and almost eliminated background noise. The timing of oral tests was done by means of an 'Interval Timer'. By a simple lever device the timing could be started at the same instant as the recording, with a soft buzzer marking off any set periods on the disk.

(iii) *The population and sample*

For administrative reasons it was decided to use the children of the north-eastern division of Surrey—within an age range of 13+ to just under 15 years—as a resident

population from which the sample could be drawn. In Surrey, under the age of 15 years, approximately 22% of the pupils are in grammar schools, 2% are in technical and similar schools, and 76% are in modern schools.[1] We decided to draw a sample from the population in such a way that these percentages were accurately represented. In addition, 429 children, from eight different schools, were given the Thurstone primary mental abilities test so that an estimate could be made of the mean ability level and standard deviation of the population. From this sample 100 children—50 of each sex—were drawn so that the correct grammar, technical and modern school quotas were retained, and so that the means and standard deviations represented those estimated for the resident population.[2]

Our rather elaborate quota-sampling procedure was adopted for a specific reason. Much of the American factorial work is based on samples drawn from the College or University levels of the population. In situations of this kind, Vernon (1950) has shown that student selection is based on 'a great variety of socio-economic, educational and other influences which are unlikely to be correlated with any factor other than *g*. Hence the *g*-variance only is reduced and no great harm is done.'

But where a representative sample of children is used, as in Blakey's (1940) re-analysis of the Brown-Stephenson (1933) data, *g* is much larger than all the other factors combined. It should be noted that although Blakey tried to rotate his factors into simple structure, a meaningful solution could not be obtained without retaining a general factor accounting for 41·2% of the total variance. In fact, Blakey's re-analysis strongly supports Vernon's (1950) contention that 'in unselected populations of adults or children...it is psychologically foolish as well as mathematically difficult to belittle it (*g*)'.

As the results of this investigation are to be interpreted from a hierarchical point of view, it is not our aim to reduce the size of *g* by either restricting the sample to a particular ability level or by rotating its variance on to the group factors—as in true simple structure solutions. Consequently, a representative sample of children is fundamental to the experimental design.

(iv) *The battery of measures*[3]

A brief description is given below of the twenty-six variables selected for the investigation.[4] The names used are those most frequently encountered in the literature. In general, the tests fall into three sections, namely: (1) the instructions to the subject with a few examples; (2) some fore-exercises in which the subject is given the opportunity to try a task within fairly flexible time limits; and (3) the test proper. Time limits and scoring procedures have been summarized in Table 1. With the short test-descriptions we will indicate the principal articles in which the same or a similar variable has appeared:

(1*w*) Handwriting—copying a familiar passage at a natural rate of writing (Cattell, 1933*a*, *b*).
(2*w*) Rhyming words—supplying words to rhyme with 'bake', 'chalk', etc. (Thurstone, 1941*a*).

[1] Data kindly supplied by the County Inspector, Dr F. S. Marston.

[2] In the final sample, the means and standard deviations were as follows: girls 103·4 and 16·45; boys 100·6 and 15·85; combined 102 and 16·2. For a fuller description see Rogers (1952*b*, *c*).

[3] Microfilm copies of the tests used by Thurstone (1936, 1937, 1940*b*, 1941*b*) and Carroll (1941*a*) were obtained from the American Documentation Institute, 1719 N. Street, N.W. Washington, 6, D.C.

[4] The variables were selected from the 60-odd fluency tests that have been described in the literature. A trial experiment was run on a group of technical school boys, and this helped to modify the test material—where necessary—and determine the time-limits for each test. In the trial run the recording techniques were stabilised and the reactions of the subjects to this novel procedure were determined by questionnaire and discussion. It was not difficult to establish rapport because of the intrinsic interest of the experiment.

(3*w*) Similes—completing in three ways statements such as 'The native boy swam through the water like...' (Stumberg, 1928; Rogers, 1952*c*).
(4*w*) Given letters—producing words containing the letters A and E (Thurstone, 1941*a*; Taylor, 1947).
(5*w*) Unfinished story—writing more to an unfinished story (Hargreaves, 1927).
(6*w*) Grammar—detecting and correcting the grammatical errors in a number of sentences (Thurstone, 1938).
(7*w*) Completion—completing sentences like that following: 'A contest of speed is a (r)....' (Thurstone, 1938).
(8*w*) Things to eat—furnishing the names of things that can be eaten (Cattell, 1936; Eysenck, 1947).
(9*w*) Spelling—differentiating the correct from the incorrect words in a list of 100 (Thurstone, 1938).

Table 1. *The battery of twenty-six measures**

Code no.	Test	Time limits (min.) Fore-exercise	Time limits (min.) Test proper	Scoring formula
1*w*	Handwriting	3	4½	(Total no. of words)
2*w*	Rhyming words	3	4	R
3*w*	Similes	4	3	R
4*w*	Given letters	6	6	R
5*w*	Unfinished story	2½	6	R (total no. of significant words)
6*w*	Grammar	2½	7	R
7*w*	Completion	3½	4½	R
8*w*	Things to eat	2½	4	R (Repetitions not accepted)
9*w*	Spelling	3	5	R −W
10*w*	Disarranged words	2½	4	R
11*w*	Birds	3	6	R
12*w*	Inventive opposites	4	5	R
13*w*	Letter star	5	10	(No. of words in correct items)
14*w*	Controlled associations	5	6	R
15*w*	Number patterns	8	5	R
16*w*	Indirect fluency	—	—	(Total no. of words)
17*w*	Verbal meaning	7	4	R
18*w*	Reasoning	8	6	R
19*o*	Form naming	2½	1	R
20*o*	Pictures	1½, 1	2	R (Total no. of ideas)
21*o*	First letters	1½	2½	R
22*o*	Topic: parcel	2	2	R (No. of relevant words)
23*o*	Inkblots	1½, 1, ½	2¾	R (No. of ideas)
24*o*	Animals	1½	2½	R
25*o*	Picture description	3½	2	R (No. of relevant words)
26*o*	Natural rate of reading	1½	¾	(Total no. of words)

* *w* signifies a written test; *o* signifies an oral, recorded test.

(10*w*) Disarranged words—unscrambling letters such as 'ebar' so that they spell the name of an animal (Thurstone, 1938; Carroll, 1941*a*).
(11*w*) Birds—listing the names of birds within a given time limit (Eysenck, 1947).
(12*w*) Inventive opposites—supplying antonyms to words like 'narrow', 'large', etc. (Thurstone, 1938).
(13*w*) Letter star—inventing meaningful arrangements of words to represent patterns such as *Y*S, in which the asterisk can stand for any word but the second and fourth must commence with the letters indicated (Carroll, 1938).
(14*w*) Controlled association—writing words similar in meaning to 'quiet', etc. (Whipple, 1915; Thurstone, 1938).
(15*w*) Number patterns—calculating the integer which is required to link a number series (Thurstone, 1941*b*).
(16*w*) Indirect fluency—total number of words used in answering the questions given after test (26*o*) (Rogers, 1952*c*).
(17*w*) Verbal meaning[1]—checking the synonym—in a list of four—of words such as moist, quick, etc. (Thurstone, 1947).

[1] Inserted as reference test.

(18*w*) Reasoning[1]—educing the next letter in a series such as 'aabccdeefgg...' (Thurstone, 1947).
(19*o*) Form naming—naming the geometrical figures (stars, circles, squares, crosses, triangles), printed in random order on a chart (Woodworth & Wells, 1911; Carroll, 1941*a*).
(20*o*) Pictures—giving responses to pictures of a tree and a street corner (Cattell, 1936).
(21*o*) First letters—producing words that begin with the letter M (Thurstone, 1947).
(22*o*) Topic: Parcel—making up a story about a parcel (Cattell, 1936).
(23*o*) Inkblots—stating the things which can be seen in three different inkblots (Cattell, 1936).
(24*o*) Animals—saying the names of as many different animals as possible (Eysenck, 1947).
(25*o*) Picture description—making up a story to describe a picture the subject has been shown (Carroll, 1941*a*; Rogers, 1952*c*).
(26*o*) Natural rate of reading—reading at a natural rate a passage taken from the Binet scale (Studman, 1934; Rogers, 1952*c*).

(v) *Reliability and validity*[2]

Special problems occur when one attempts to estimate the reliability of an associational or fluency type of test. For instance, Guilford (1950) draws a distinction between speed and power tests, and maintains that with the former any internal-consistency estimate of r_{tt} would be mis-applied. However, Notcutt (1943) and Hargreaves (1927) have shown that fluency tests correlate as highly with intelligence as with speed, so Guilford's objection is partly met.

A further difficulty originates from the method of scoring fluency tests. Where all, or nearly all, responses to a stimulus are acceptable, then the split-half coefficient of reliability would be plus one. But in the cases where answers can be correct or incorrect, it would appear that such an estimation is more defensible.

Split-half reliabilities were calculated for the twelve tests where it seemed applicable. These ranged from 0·71 to 0·99 with a mean of 0·88. In addition, minimum indices of reliability were calculated from the communalities of the test battery, cf. Guilford (1950). These ranged from 0·56 to 0·90 with an average of 0·80. Because factorial estimates represent minimum rather than actual values, the reliability of the battery as a whole was taken to be adequate.

The validity of the battery was also determined by factorial procedures. Of the twenty-six tests employed in the study, only six had validities below 60%, while eleven were higher than 70%. Our results compare favourably with other studies in which hypotheses are erected on the basis of common-factor variance at about 50–60% level—cf. Guilford (1946).

III. A factor analysis of the data

Thurstone's centroid method of factor analysis—a modification of the Burt simple summation technique—was chosen for this study. Thurstone uses it to obtain a multiple factor solution, but Vernon (1946, 1947*a*, *b*) has shown how it can be employed to arrive at a general group factor pattern.

The original inter-test correlations are shown in Table 2. These were factor analysed until five factors had been extracted. The unrotated factors are shown in Table 3.

(i) *The extraction of common factors*

A problem of long standing in factor analysis concerns the number of factors which should be extracted from a matrix of unknown rank. Vernon (1949) has discussed about

[1] Inserted as reference test.

[2] For a statistical treatment of these topics the reader is referred to Rogers's (1952*c*) thesis.

Table 2. *Matrix of inter-test correlations*

	1*w*	2*w*	3*w*	4*w*	5*w*	6*w*	7*w*	8*w*	9*w*	10*w*	11*w*	12*w*	13*w*	14*w*	15*w*	16*w*	17*w*	18*w*	19*o*	20*o*	21*o*	22*o*	23*o*	24*o*	25*o*	26*o*
1*w*	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2*w*	665	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
3*w*	630	645	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
4*w*	485	730	450	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
5*w*	790	688	655	490	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
6*w*	633	750	595	503	595	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
7*w*	463	613	518	558	428	648	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
8*w*	613	718	645	730	710	675	585	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
9*w*	185	405	200	398	210	478	403	380	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
10*w*	510	645	468	670	578	623	560	703	455	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
11*w*	380	578	480	500	398	570	678	613	333	640	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
12*w*	623	743	570	618	590	678	765	685	358	640	625	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
13*w*	698	738	645	603	748	678	495	735	275	605	470	613	—	—	—	—	—	—	—	—	—	—	—	—	—	—
14*w*	553	618	538	595	543	550	558	650	245	668	635	658	555	—	—	—	—	—	—	—	—	—	—	—	—	—
15*w*	300	343	293	323	233	398	263	400	225	485	403	330	433	375	—	—	—	—	—	—	—	—	—	—	—	—
16*w*	105	133	030	248	098	055	−068	245	095	213	100	−055	195	153	163	—	—	—	—	—	—	—	—	—	—	—
17*w*	575	798	618	670	565	713	778	713	398	623	630	795	683	590	348	053	—	—	—	—	—	—	—	—	—	—
18*w*	465	650	423	578	438	655	573	545	485	585	495	663	583	450	423	123	673	—	—	—	—	—	—	—	—	—
19*o*	448	525	300	535	365	415	423	603	160	515	548	508	553	453	338	210	505	408	—	—	—	—	—	—	—	—
20*o*	315	158	298	160	198	255	390	263	−033	180	310	413	308	228	008	−088	355	258	363	—	—	—	—	—	—	—
21*o*	478	643	500	548	423	440	470	558	263	430	525	593	528	548	198	190	533	450	390	413	—	—	—	—	—	—
22*o*	300	210	195	193	233	168	400	193	015	155	258	420	295	258	073	158	365	233	228	653	350	—	—	—	—	—
23*o*	460	348	373	223	355	305	438	308	120	280	320	483	360	388	053	−070	423	305	273	660	490	658	—	—	—	—
24*o*	205	345	350	185	198	193	430	303	120	223	430	388	325	350	108	028	325	298	310	505	583	388	475	—	—	—
25*o*	343	328	228	263	308	215	283	293	028	235	245	335	360	220	038	078	335	250	228	618	448	643	565	335	—	—
26*o*	588	688	553	605	588	610	645	693	345	570	625	675	560	678	223	175	718	563	530	235	530	333	098	388	275	

Correlations of 0·197 and 0·256 are significant at the 5% and 1% levels of significance respectively.
Decimal points omitted.

twenty of the possible criteria, mostly based on empirical evidence, and the ones which we have used are listed below:

(1) Guilford & Lacey (1947)—four factors significant.
(2) Burt & Banks (1946–7)—two factors significant.
(3) Tucker (cf. Blakey, 1940)—five factors significant; possibly more.
(4) Coombs (1941)—five factors significant; possibly more.
(5) McNemar (1942)—five factors significant.

Table 3. *Unrotated factor matrix*

	Factor loadings				
Test	k_1	k_2	k_3	k_4	k_5
1*w*	738	102	−420	−015	101
2*w*	849	230	−042	112	036
3*w*	694	082	−245	−032	231
4*w*	737	238	146	−005	−195
5*w*	715	206	−470	−060	097
6*w*	770	242	−107	292	105
7*w*	765	−098	205	292	216
8*w*	837	270	−043	−151	−002
9*w*	412	200	220	315	−223
10*w*	759	292	135	−026	−135
11*w*	730	027	325	−076	172
12*w*	849	−065	092	222	154
13*w*	807	186	−279	−077	−076
14*w*	746	107	114	−152	097
15*w*	417	323	117	−062	−127
16*w*	146	264	074	−361	−295
17*w*	854	055	046	259	075
18*w*	717	124	153	257	−166
19*o*	629	072	095	−216	011
20*o*	473	−648	−114	−113	016
21*o*	712	−205	096	−186	−035
22*o*	443	−645	−127	053	−056
23*o*	547	−581	−222	099	−136
24*o*	490	−415	188	−190	095
25*o*	477	−507	−218	−115	−276
26*o*	773	189	145	−061	311

Decimal points omitted.

Because of the divergent evidence the allocation of significance to a particular factor must be attended with caution. However, the following assumptions would appear to be reasonable. Factors one and two are certainly significant, and it seems likely that factor three is as well. Factors four and five are more doubtful, but they have been retained to provide rotational freedom and for tentative description.

(ii) *The rotational problem*

Inasmuch as we are dealing with cognitive abilities, it was decided to rotate the arbitrary reference frame in such a way that all the factors became predominantly positive, and possessed the greatest amount of psychological meaningfulness. Eight orthogonal rotations were necessary to stabilize the five factors of the matrix—cf. Rogers (1952*c*). Throughout the rotations certain consistent principles have been followed and these are outlined below:

(*a*) *Alexander's* (1935) *method of rotation* consists of placing known reference tests in a battery and rotating through them so that the position of related variables is clearly

established. Reference test (17*w*) was used in this manner to help stabilize the general factor.

(*b*) *Eysenck's* (1950) *criterion analysis* employs an external criterion to arrive at a unique, invariant solution. We used the Thurstone primary mental abilities test as a measure of general ability, and rotated the first factor of the solution into maximum correlation (*r* 0·91) with it.

Table 4. *Rotated factor matrix*

Test	Rotated factor loadings I	II	III	IV	V	h^2 (unrotated)	h^2 (rotated)
1*w*	568	147	616	083	087	741	739
2*w*	819	−022	309	135	065	790	788
3*w*	568	079	476	−011	218	603	604
4*w*	716	−039	084	368	045	659	659
5*w*	533	061	691	127	081	787	788
6*w*	800	−051	332	−037	−063	759	758
7*w*	830	150	−009	−186	145	769	768
8*w*	721	−051	351	322	221	799	799
9*w*	551	−089	−136	135	−242	407	407
10*w*	735	−098	130	350	086	698	697
11*w*	715	011	−043	088	391	676	674
12*w*	855	182	121	−081	144	806	807
13*w*	645	097	514	288	072	776	777
14*w*	660	013	169	198	335	613	615
15*w*	418	−199	071	303	034	312	313
16*w*	046	−150	041	528	096	313	315
17*w*	875	106	172	−015	037	807	808
18*w*	778	050	007	164	−109	646	647
19*o*	520	046	138	266	309	457	457
20*o*	265	717	087	−019	280	670	670
21*o*	564	328	076	232	329	594	594
22*o*	293	728	049	−066	108	634	634
23*o*	383	738	153	001	012	714	715
24*o*	363	412	−090	046	423	492	491
25*o*	250	690	156	233	062	622	621
26*o*	743	−119	196	014	388	756	755

Decimal points omitted.

(*c*) *Simple structure.* An inspection of the rotated factor matrix shows that, apart from the general factor which has been retained, Thurstone's (1940*a*, 1947) requirements for simple structure are met—also cf. Blakey (1940).

(*d*) *Simultaneous hierarchical structure* is a modification of Cattell's (1946) principle of 'proportional profiles'.[1] It consisted of solving this matrix, along with others also designed to study verbal fluency, so that essentially the same cognitive factor plan emerged from them all.

(iii) *Interpretation of the factors*

In accordance with Thurstone's (1938) procedure a projection is regarded as zero if it is within the range of plus or minus 0·20 and, for the purpose of operationally defining a factor, greatest weight has been given to those projections which are larger than 0·40. The plan of significant loadings (Table 5) will be examined by columns.

[1] Strictly speaking, Cattell (1944, 1952) reserves the concept of 'proportional profiles' for simple structure solutions only, and in this sense we do not follow his lead. He argues that the scientific principle of parsimony is best applied in factor analysis 'with respect to a whole population of factor analyses rather than a single matrix'. In other words, his aim is to arrive at a factorial solution which fits in with the simple structure solutions of many experiments. In this investigation we have adopted the principle of rotating to what may be termed 'simultaneous hierarchical structure', rather than 'simultaneous simple structure'.

Factor I. The first factor has significant loadings on twenty-five of the tests, the majority of them being above 0·40. The ten tests with the highest projections are as follows:

(17*w*) Verbal meaning	0·88	(18*w*) Reasoning	0·78
(12*w*) Inventive opposites	0·86	(10*w*) Disarranged words	0·74
(7*w*) Completion	0·83	(26*o*) Natural rate of reading	0·74
(2*w*) Rhyming words	0·82	(4*w*) Given letters	0·72
(6*w*) Grammar	0·80	(8*w*) Things to eat	0·72

Table 5. *Plan of the significant loadings*

		Significant factor loadings				
Code no.	Test	I	II	III	IV	V
1*w*	Handwriting	57	—	62	—	—
2*w*	Rhyming words	82	—	31	—	—
3*w*	Similes	57	—	48	—	22
4*w*	Given letters	72	—	—	37	—
5*w*	Unfinished story	53	—	69	—	—
6*w*	Grammar	80	—	33	—	—
7*w*	Completion	83	—	—	—	—
8*w*	Things to eat	72	—	35	32	22
9*w*	Spelling	55	—	—	—	−24
10*w*	Disarranged words	74	—	—	35	—
11*w*	Birds	72	—	—	—	39
12*w*	Inventive opposites	86	—	—	—	—
13*w*	Letter star	65	—	51	29	—
14*w*	Controlled association	66	—	—	—	34
15*w*	Number patterns	42	—	—	30	—
16*w*	Indirect fluency	—	—	—	53	—
17*w*	Verbal meaning	88	—	—	—	—
18*w*	Reasoning	78	—	—	—	—
19*o*	Form naming	52	—	—	27	31
20*o*	Pictures	27	72	—	—	28
21*o*	First letters	56	33	—	23	33
22*o*	Topic: parcel	29	73	—	—	—
23*o*	Inkblots	38	74	—	—	—
24*o*	Animals	36	41	—	—	42
25*o*	Picture description	25	69	—	23	—
26*o*	Natural rate of reading	74	—	—	—	39
	Variance (%)	38·8	9·8	7·4	4·5	4·4

Decimal points omitted.

It requires little inspection of these tests to deduce that they are all of a verbal or educational nature. Consequently, an operational definition of the factor would be 'general verbal ability'. As the original correlation matrix was essentially positive, and as variants of the reference test (17*w*) have been used in numerous batteries as measures of 'general ability', it can be safely assumed that the factor depends in part on the more fundamental *g*. However, without further analysis the test loadings on *g* cannot be determined.

The evidence indicates that factor I is composed of *g* plus the *v* : *ed* factor described by Vernon (1950). Actually, some distinction should be made between 'mathematical *g*' and 'psychological *g*'. Factor I, being the first general factor, can rightly be referred to as 'mathematical *g*', but on the basis of this experiment, it cannot be identified with 'psychological *g*' in the strict Spearman (1927) sense.

Most of the tests on factor I have been previously used as measures of some kind of fluency. On analysis it becomes clear that the major part of their variance can be accounted for by the $g+v$: *ed* factors, and the dichotomy into *V* and *W* which Thurstone finds, is not evident. 'General verbal ability' has a total variance of 38·8%, substantially more than all the other factors combined.

Factor II has significant loadings on six of the oral tests. They are:

(23*o*) Inkblots	0·74	(25*o*) Picture description	0·69
(22*o*) Topic: parcel	0·73	(24*o*) Animals	0·41
(20*o*) Pictures	0·72	(21*o*) First letters	0·33

An inspection of the above variables suggests that this factor is one of oral facility, or ability to give spontaneous oral expression to one's ideas in an effective and coherent manner. It is quite noteworthy that form naming (19*o*) and natural rate of reading (26*o*) which do not depend on spontaneous oral expression, do not have significant loadings on oral facility. Where the test requires the manipulation of responses and the expression of them orally, it has loadings on the factor of oral facility.

It can be seen that first letters (21*o*), which Thurstone uses as a word-fluency test, has its largest loading on the general factor followed by a small saturation on oral facility. Thurstone's multiple factor solution does not allow it any variance on a general factor.

Factor II seems to substantiate Carroll's (1941*a*) 'facility and fluency in oral speech'. It has been suggested that this factor may represent the subject's ease and confidence in a novel testing situation. However, Carroll's (1941*a*) results indicate that this is unlikely, and the fact that neither the first (19*o*) nor the last (26*o*) of the recorded tests have loadings on oral facility, also leads to a rejection of the hypothesis.

The second factor accounts for 9·8% of the total variance.

Factor III has significant projections on seven of the variables, i.e.:

(5*w*) Unfinished story	0·69	(8*w*) Things to eat	0·35
(1*w*) Handwriting	0·62	(6*w*) Grammar	0·33
(13*w*) Letter star	0·51	(2*w*) Rhyming words	0·31
(3*w*) Similes	0·48		

A study of the operations essential in these tests reveals that the factor involves a 'facility in writing'. It seems to represent a fluency in the written sphere, just as factor II represents a fluency in the oral sphere. With general verbal ability held constant, it is significant that oral facility and facility in writing do not overlap at any point.

The interpretation of tests (5*w*), (1*w*) and (13*w*) seems quite straightforward. The test material revealed that a high—or low—score was a function of the speed at which the subject could write. Unfinished story (5*w*), with a loading of 0·69, was inserted in the battery to parallel (22*o*)—topic: parcel. Its saturation on factor III emphasizes an essential difference between the facility called forth in written as distinct from oral situations.

Similes (3*w*), which was usually answered by means of a phrase or incomplete sentence, also depended on facility in writing, as did things to eat (8*w*) and rhyming words (2*w*). Grammar (6*w*) was more difficult to interpret in this way, and it is perhaps significant that its projection on the factor vanished in a subsequent analysis—cf. Rogers (1952*c*).

The amount accomplished in each of these tests, with the possible exception of grammar, clearly depended on the subject's facility or fluency in writing. The test situations varied considerably but they all appeared to have this in common. Other variables which required written responses were also considered, but in no case did it appear that facility in writing would affect the result.

Variance on factor III accounts for 7·4% of the total.

Factor IV. The tests which have loadings above 0·20 on this factor are:

(16*w*) Indirect fluency	0·53	(13*w*) Letter star	0·29
(4*w*) Given letters	0·37	(19*o*) Form naming	0·27
(10*w*) Disarranged words	0·35	(21*o*) First letters	0·23
(8*w*) Things to eat	0·32	(25*o*) Picture description	0·23
(15*w*) Number patterns	0·30		

As this factor has only one test (16*w*) with a loading above 0·40, any interpretation of it must be attended with caution. In looking for an underlying unity in the variables listed above, it seemed that factor IV might involve a fluency in dealing with words in which one or more formal restrictions were placed on the response, but little reference was made to its meaning. This interpretation is somewhere near that of Thurstone (1947), Taylor (1947), and would account for the loadings on given letters (4*w*), disarranged words (10*w*), letter star (13*w*), form naming (19*o*), first letters (21*o*) and possibly things to eat (8*w*).

Factor V was non-significant on the criteria of Burt and Guilford. Projections above 0·20 are as follows:

(24*o*) Animals	0·42	(19*o*) Form naming	0·31
(11*w*) Birds	0·39	(20*o*) Pictures	0·28
(26*o*) Natural rate of reading	0·39	(3*w*) Similes	0·22
(14*w*) Controlled association	0·34	(8*w*) Things to eat	0·22
(21*o*) First letters	0·33	(9*w*) Spelling	−0·24

The nature of this factor, like that of factor IV, is rather obscure. It has been tentatively identified as an ideational fluency, in which the subject expresses ideas by the use of words and their meanings. This is similar to Taylor's (1947) identification, with quantity not quality being the most important feature of the test.

A further analysis revealed that this factor was relatively unstable, and of little value for indicating relations with orectic dimensions. It has been retained because other researches show that the hypothesis is tenable.

IV. Hypotheses and conclusions

It is now possible to advance hypotheses in answer to the problems outlined at the beginning of the article. All hypotheses are relative to this sample and battery of tests.

(1) The major variance of a battery of fluency tests can be accounted for by a general factor. On the basis of other research evidence, it seems clear that this is *g* plus the *v* : *ed* factor described by Vernon. The factor of 'general verbal ability', as it has been called, accounts for more variance than all the other factors combined.

(2) Two distinct fluency factors can be identified. The first, with a variance of 9·8%, has been defined as an 'oral facility', and the second, with a variance of 7·4%, has been established as a 'facility in writing'. It would appear that these orthogonal factors have often been confused in previous studies.

(3) The most parsimonious interpretation of fluency would be as follows:

$f = g + v$: *ed* + *either* 'oral facility' *or* 'facility in writing'. Two other factors have been tentatively identified as a 'fluency in dealing with words' and 'ideational fluency' respectively. They straddle the oral and written fields and also overlap with each other on some tests.

(4) A hierarchical interpretation of abilities works on the law of parsimony described by William of Occam. It is a scientific maxim to prefer the least complex of alternative

explanations of phenomena. This being so, it seems fruitless to extract large numbers of factors from a matrix and then build up their variance by rotating off the most important factor, namely, the general one.

The investigation re-emphasizes the importance of the $g + v$: *ed* factors found in so many British researches, and suggests that two other verbal factors may have some importance in education. It is a reasonable deduction that, where time limits are imposed on written and oral tests, the operation of group factors in these two fields would influence individual performance.

Finally, because of the ease with which fluency tests can be given and the simplicity of scoring them, we suggest that a small battery of say three measures, could be tried as part of the selection process at 11 plus. Their g content is almost as high as the best v tests and they are much easier to apply.

REFERENCES

ALEXANDER, W. P. (1935). Intelligence, concrete and abstract. *Brit. J. Psychol. Monogr. Suppl.* XIX, pp. 177.

BECHTHOLDT, H. P. (1947). Factorial investigation of the perceptual-speed factor. *Amer. Psychol.* II, 304–5.

BLAKEY, R. (1940). A re-analysis of a test of the theory of two factors. *Psychometrika*, V, 121–36.

BROWN, W. & STEPHENSON, W. (1933). A test of the theory of two factors. *Brit. J. Psychol.* XXIII, 352–70.

BURT, C. (1917). *The Distribution and Relations of Educational Abilities*, p. 93. London: King.

BURT, C. & BANKS, C. (1946–7). A factor analysis of body measurements for British adult males. *Ann. Eugen. Lond.*, XIII, 238–56.

CAREY, N. (1916). Factors in the mental processes of school children. III. Factors concerned in the school subjects. *Brit. J. Psychol.* VIII, 170–82.

CARROLL, J. B. (1938). Diversity of vocabulary and the harmonic series law of word-frequency distribution. *Psychol. Rec.* II, 379–86.

CARROLL, J. B. (1941*a*). A factor analysis of verbal abilities. *Psychometrika*, VI, 279–307.

CARROLL, J. B. (1941*b*). A factor analysis of verbal abilities. Washington, D.C.: American Documentation Institute. (*Doc.* 1597), p. 58.

CATTELL, R. B. (1933*a*). Temperament tests. I. Temperament. *Brit. J. Psychol.* XXIII, 308–29.

CATTELL, R. B. (1933*b*). Temperament tests. II. Tests. *Brit. J. Psychol.* XXIV, 20–49.

CATTELL, R. B. (1936, 1948). *A Guide to Mental Testing*, p. 312. London: University of London Press.

CATTELL, R. B. (1944). 'Parallel proportional profiles' and other principles for determining the choice of factors by rotation. *Psychometrika*, IX, 267–83.

CATTELL, R. B. (1946). *Description and Measurement of Personality* p. 602. New York: World Book Co.

CATTELL, R. B. (1952). *Factor Analysis*, p. 462. New York: Harper.

COOMBS, C. H. (1941). A criterion for significant common factor variance. *Psychometrika*, VI, 267–72.

CURETON, E. E. (1947). The verbal relations factor and vocabulary. *Amer. Psychol.* II, 286–7.

DAVEY, C. M. (1926). A comparison of group verbal and pictorial tests of intelligence. *Brit. J. Psychol.* XVII, 27–48.

EYSENCK, H. J. (1939). Critical review of 'primary mental abilities' by L. L. Thurstone. *Brit. J. Educ. Psychol.* IX, 270–5.

EYSENCK, H. J. (1947). *Dimensions of Personality*, p. 308. London: Routledge and Kegan Paul.

EYSENCK, H. J. (1950). Criterion analysis. *Psychol. Rev.* LVII, 38–53.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development*, pp. 261. London: Dent.

GUILFORD, J. P. (1946). New standards for test evaluation. *Educ. Psychol. Meas.* VI, 427–38.

GUILFORD, J. P. (1950). *Fundamental Statistics in Psychology and Education*, p. 633. New York: McGraw-Hill.

GUILFORD, J. P. & LACEY, J. L. (1947). Printed classification tests. *Army Air Forces Aviat. Psychol. Prog. Res. Rep.* no. 5, pp. 919. Washington, D.C.: U.S. Gov. Print. Office.

HARGREAVES, H. (1927). The 'faculty' of imagination. *Suppl. Brit. J. Psychol. Monogr.* X, p. 74.

HOLZINGER, K. J. & HARMAN, H. H. (1938). Comparison of two factorial analyses. *Psychometrika*, III, 45–60.

KELLEY, T. L. (1947). *Fundamentals of Statistics*, pp. 755. Cambridge: Harvard University Press.
MCNEMAR, Q. (1942). On the number of factors. *Psychometrika*, VII, 9–18.
NOTCUTT, B. (1943). Perseveration and fluency. *Brit. J. Psychol.* XXXIII, 200–8.
ROGERS, C. A. (1952*a*). The orectic relations of verbal fluency. (To appear.)
ROGERS, C. A. (1952*b*). The orectic relations of mathematically derived fluency scores. (To appear.)
ROGERS, C. A. (1952*c*). A factorial study of verbal fluency and related dimensions of personality, p. 364. University of London Ph.D. thesis.
SPEARMAN, C. E. (1904). 'General intelligence', objectively determined and measured. *Amer. J. Physiol.* XV, 201–93.
SPEARMAN, C. E. (1927). *The Abilities of Man*, pp. 415. London: Macmillan.
STUDMAN, G. L. (1934). Spearman factors and psychiatry. VI. Measurement of the speed and flow of mental activity. *Brit. J. Med. Psychol.* XIV, 124–31.
STUMBERG, D. (1928). A study of poetic talent. *J. Exp. Psychol.* XI, 219–34.
TAYLOR, C. W. (1947). A factorial study of fluency in writing. *Psychometrika*, XII, 239–62.
THURSTONE, L. L. (1936). Psychological tests used in a factorial study at the Lane Technical High School in Chicago, Illinois, 1936. Washington, D.C.: American Documentation Institute. (*Doc.* 1337), pp. 178.
THURSTONE, L. L. (1937). Psychological tests used in a factorial study at the Hyde Park High School, Chicago, Illinois, 1937. Washington, D.C.: American Documentation Institute. (*Doc.* 1329), pp. 140.
THURSTONE, L. L. (1938). Primary mental abilities. *Psychometr. Monogr.* I, pp. 121.
THURSTONE, L. L. (1940*a*). Experimental study of simple structure. *Psychometrika*, V, 153–68.
THURSTONE, L. L. (1940*b*). Psychological tests used in a study of mental abilities. Washington, D.C.: American Documentation Institute. (*Doc.* 1317), pp. 304.
THURSTONE, L. L. (1941*a*). Factorial studies of intelligence. *Psychometr. Monogr.* II, p. 94.
THURSTONE, L. L. (1941*b*). Psychological tests used in a factorial study of eighth-grade children. Washington, D.C.: American Documentation Institute. (*Doc.* 1434), pp. 188.
THURSTONE, L. L. & THURSTONE, T. G. (1947). SRA primary mental abilities. Ages 11–17, form AH. Chicago: Science Research Associates, pp. 11.
THURSTONE, L. L. (1947). *Multiple-factor Analysis*, pp. 535. Chicago: University of Chicago Press.
VERNON, P. E. (1946). Statistical methods in the selection of Navy and Army personnel. *Suppl. J. R. Statist. Soc.* VIII, 139–53.
VERNON, P. E. (1947*a*). Research on personnel selection in the Royal Navy and the British Army. *Amer. Psychol.* II, 35–51.
VERNON, P. E. (1947*b*). Psychological tests in the Royal Navy, Army and A.T.S. *Occup. Psychol.* XXI, 53–74.
VERNON, P. E. (1949). How many factors? London: *Admiralty Report*, p. 10.
VERNON, P. E. (1950). *The structure of human abilities*, pp. 160. London: Methuen.
VERNON, P. E. & PARRY, J. B. (1949). *Personnel selection in the British forces*, pp. 324. London: University of London Press.
WHIPPLE, G. M. (1915). *Manual of Mental and Physical Tests. Part II. Complex Processes*, pp. 336. Baltimore: Warwick and York.
WOODROW, H. (1939). The common factors in fifty-two mental tests. *Psychometrika*, IV, 99–108.
WOODWORTH, R. S. & WELLS, F. L. (1911). Association tests. *Psychol. Monogr.* XIII, no. 57, pp. 85.

(*Manuscript received* 19 *November* 1952)

FIVE YEARS IN THE OXFORD CHAIR

By GEORGE HUMPHREY

Institute of Experimental Psychology, University of Oxford

In 1947 the University of Oxford decided, after considerable debate, to found a Chair of Psychology. There had already been an Institute of Experimental Psychology which had been initiated and directed by the late Dr William Brown and his associate, William Stephenson (a brigadier in the last War); but it was able to grant only a Diploma in Psychology and advanced degrees. It took no part in the undergraduate instruction of the University. The Chair was established at the same time as a new degree course for undergraduates, including psychology as one of its subjects.

In certain ways the Honours School of Psychology, Philosophy and Physiology is unique, at least at this University, in that it forms a special kind of meeting place for the man of the laboratory and the 'humanist'. It is, perhaps, particularly over examinations that the clash of personality and training comes out. The Board of Examiners consists of two psychologists, two physiologists (one external in each), and two philosophers. The Board, as a whole, is responsible for the class in which a man is placed. Philosophers, psychologists and physiologists sit round the same table discussing examination papers in 'Psychology, and either Philosophy or Physiology' and, after the papers have been written, examination results. A physiologist or a philosopher may ask to read a psychological paper; in fact, at the wish of one of the physiologists, it has become the custom for the special paper on physiological psychology to be read by the physiologists. A philosopher or a physiologist may turn up at the practical examinations or the *viva voce* in psychology. It is not breaking any professional confidence to say that on one occasion a physiologist thought that the psychologist had been too hard in assessing a paper, and was told that what had been written was fairly routine work for a psychologist. This interchange of ideas and scrutiny of psychologists' work by philosophers and physiologists—the converse also happens at times—helps to give a special flavour to what has become known as P.P.P. The Board of Studies for Psychology, which has general control over examinations for the School and certain other matters, at present includes three philosophers, four psychologists, the reader in mental philosophy (a liaison man!), a zoologist, two physiologists, a neurologist and a head of a college who is a well-known philosopher. Here again Science and the Humanities sit round the same table to discuss the problems of the School and the Institute. Experience up to the present has shown that each side has its own distinctive and valuable contribution to make.

This organization has led to interchange of views not only about particular candidates during *vivas* and otherwise, but also on points of educational policy. The conflux of disciplines, with psychology as the middle stream, has raised in a more acute form than elsewhere the difficult problem of the relation of the Arts to the Sciences; for, of course, the same problem exists at other universities in the country and for that matter throughout the world. Here, psychology has proved itself a natural interpreter.

Oxford is the home of professional philosophers. It is safe to say that there are more of them to the square mile than anywhere else, perhaps with the exception of the Vatican.

They have at times had strong views about psychology, particularly experimental psychology. An eminent tutor used to set his pupils an essay in which they were expected to prove that experimental psychology was impossible; and he was not alone at the time in his belief. In fact, when experimental psychology was born with Fechner's *Psychophysics* in 1860, the dictum 'This is all nonsense!' seems to have arisen simultaneously from twenty Oxford colleges. One may guess that it was Freud who originally both inhibited and electrified the notion that something more should be done about psychology in this University. For in 1945 the idea of an experimental study of the mind was still a bit of a novelty in this home of ancient philosophical traditions. All this did not make things easier for a new professor, who found himself the jetsam of an acrimonious debate as to whether he should exist at all. He heard the lions on the shore sniffing round him, waiting for him to make a wrong move. Occasionally he still hears a sniff or two.

Now Oxford philosophy has a long history, a long memory and a strong arm. It reaches its conclusions patiently and carefully; they are honest conclusions, honestly reached. Though rebels and initiators exist, there is nevertheless an important body of immensely old, immensely well-grounded and well-groomed conservatism. With respect to psychology, it must be remembered that many of the objections were originally raised against the futilities of a past that we have now grown out of, and the not so long past extravagances of an early psycho-analysis too wholeheartedly promoted by well-meaning disciples. To some extent it is therefore true that to-day's psychologist would endorse many of the objections with which he has been met. Conservative opinion has feared also that the subject might be too 'easy'. This fear we are in process of removing. It has feared also that undergraduates might take the course 'for the wrong reasons', because they were interested in the morbid, and expected to find it in psychology. One can only say that they have not.

Lastly, and perhaps most important, it has feared that psychology could not provide a discipline stringent and basic enough to contribute to that liberation of the mind on which Oxford has so long and so properly prided herself. During the first year of the chair a very distinguished Oxonian, at that time Chairman of the Board of Studies for Psychology, remarked that if we could 'educate the men' we should have little trouble. This is what we have tried to do. The endeavour has been to make people see nature—human nature—as a problem. Facts there must of course be, and techniques also; but if we can make our graduates look on the world of human beings in the spirit of research, we are satisfied. Thus the whole course is 'research minded'. The amount and the variety of advanced research going on in the Institute has steadily grown since the School of Psychology, Philosophy and Physiology was begun, and now ranges from the electrical phenomena of the eye to the psychology of economic processes. Undergraduates have been described as 'regarding each question as a problem for research'. This, we feel, is as it should be. In the matter of research, as in every other respect, the Institute has enjoyed and appreciated unusually cordial relations with the Oxford hospitals.

The Institute itself is accommodated in half of an old school house together with a 'hut' in the rear. In addition, certain outlying premises are used for research. The University has promised a new building, which will make everybody much happier and bring all staff and research students together, as befits a young and growing department. There are now some forty undergraduates and research students. Owing to the scarcity of space there is a 'ceiling' for the number of undergraduates permitted to read the School, and this also

helps a young department. According to our terms of reference, training is given only in fundamental psychology, not in applications, such as educational or industrial work. The administrative relation of the Institute to the School of P.P.P. has not yet been completely clarified. In practice the Institute works very well as the major repository of psychological work in the University.

We are trying to give the student that spirit of inquiry which the philosophers have so manfully kept alive for so many years. For the rest, Pericles once remarked that that woman is best who is least spoken of for good or ill. That we feel to be true also of a University department—at least a young one!

(*Manuscript received* 10 *June* 1953)

PUBLICATIONS RECENTLY RECEIVED

Science and Human Behavior. By B. F. Skinner. New York: Macmillan Co. 1953. Pp. 461. $4.00.

In *The Behavior of Organisms,* Skinner proposed a system for the convenient formulation of behaviour data and showed that an impressive measure of lawfulness emerged from an experimental analysis of rat lever-pressing behaviour in terms of this system. Briefly, his programme sought to describe the behaviour of the individual organism as a function of external, manipulable variables with minimum recourse to theories, hypotheses, and intervening variables of any sort—neural, mental, or conceptual. Now, fifteen years later, Skinner boldly and ingeniously analyses a wide range of human behaviour in terms of the same basic programme. As before, elicited reflex behaviour is distinguished from emitted operant behaviour (which now includes verbalization), and conditioning is dichotomized into reflex conditioning (stimulus substitution) and operant conditioning. Again, the response's frequency or rate of emission is emphasized and operant conditioning is viewed as a change in this frequency which occurs under the influence of empirically determined reinforcing conditions (law of effect). Again, the chief classes of environmental variable are the reinforcing, aversive, and discriminative stimuli and the operations responsible for the states of deprivation, satiation, and emotion. Numerous examples are cited of the control which instances of these variables, either alone or in combination, may exert over concrete behaviour items. Discussion then centres successively round the manipulation of these variables by the behaving individual himself, by other humans with whom the subject interacts in a social system, and by governmental, economic, etc., social institutions.

More avowedly than ever, the system is descriptive and incapable of generating hypotheses, e.g. the former hypothesis of 'reflex reserve' has now gone. If it is asked why the system should be a descriptive behaviourism, the answer appears to lie in Skinner's distrust of reified 'inner causes' and his conviction that too liberal a use of hypothetical constructs and intervening variables leads to the pursuit of the trivial and neglect of the important. In both practice and theory, he feels, behaviour science will ultimately profit by studying 'common sense' problems and keeping to lucid, low-level concepts. There is no means of saying, for the time being, whether this may be so or not, although most present-day psychologists find the use of higher-level constructs more satisfying. The book's primary aim seems to be to wean its reader away from the popular and traditional view of man as a free agent whose behaviour is the product of spontaneous inner changes toward a deterministic view that 'what a man does is the result of specifiable conditions and that once these conditions have been discovered, we can anticipate and to some extent determine his actions'. Accordingly, the book does not concern itself with reports of new behaviour discoveries or of contemporary psychological work; it merely presents a set of descriptive concepts and applies them to a sufficient number of familiar behaviour examples to illustrate how they should be used.

Skinner's discussions of scientific methodology aside, perhaps the most important parts of this book are those which: show that frequency of response is basic to many firmly accepted psychological concepts such as tendency, set, and habit strength; illustrate the pervasive, everyday influence of reinforcement in shaping and maintaining behaviour; provide a penetrating and provocative analysis of the effects of punishment which ought to stimulate experiment and may have far-reaching practical consequences; and, through the notion of behaviour control, supply a single framework for considering such diverse social organizations as law, education, religion, psychotherapy, and economics. For those contributions, and for the acute analyses of so many current psychological concepts, this book should be read profitably by every serious student.

I. M. L. HUNTER

History of American Psychology. By A. A. Roback. New York: Library Publishers. 1952. Pp. xiv+426. $6.00.

The many readers of Dr Roback's books in this country will approach this volume with a lively and confident hope that they will find not only a record of facts about and opinions of many famous people in American psychology, but also no lack of original comments, of attempts to sum up the significance of particular movements, and of excursions into controversial matters by the author himself. They will not be disappointed.

For most of us, probably, American psychology begins with William James, and proceeds rapidly through a number of 'schools' at war one with another to the rather bewildering diversity of the present day. It was high time that somebody should tell us more, in a readable and concise manner, about the earlier developments. This Dr Roback has now done and done well. To one reader, at any rate, the story of the beginnings, 'In the Dominion of Physics and the Empire of Theology', is in many ways the most interesting part of this entire volume. It shows a terrific and persistent search for systems from which American psychology has never really escaped. It helps the reader to understand how, from the very beginning, American psychological thought has been peculiarly open to influences from the outside, in the early days from the Scottish realists; later from the German theorists; then from the expansion of Natural Science in most of the European countries. At the same time he will realize, probably more forcibly than ever before, that the welcome which in America has been accorded to ideas and methods coming from outside America never has led, and still does not lead, to anything like imitation. Genuine independence of thought and practice are there from the start in Johnathan Edwards, and Thomas Upham, Frederick Raveh, Benjamin Rush, Noah Porter and others of whom, no doubt, most of us knew little, but now can learn more. It is intensely interesting that in these early days there was already the hankering after text-books, the deep regard for the general approach, which very often indeed was allowed to determine what was to be observed in the particular work. The theological and transcendental settings have disappeared, but, in the welter of current empiricism, the ardent search for the system, started by the pioneers, proceeds with undiminished vigour.

Roback calls part II of his History 'Psychology comes of Age'. Here he is concerned excellently with William James, then with Stanley Hall, G. T. Ladd, J. M. Cattell, Baldwin and, at somewhat greater length, with Titchener and Munsterberg. All of these are presented as persons, and not merely as contributors to psychology; but that is always a feature of this book. Some people will very likely not think highly of the Titchener section (there are one or two expressions that must appear odd to an English reader, as when the violence of the American Psychological Association is contrasted with 'the cool stratosphere of the British Association for the Advancement of Science'), but Munsterberg is extremely well and sympathetically presented.

The period from about 1900 to very recent times is given 'The Schools' as a descriptive heading. Probably this is fair enough, though it may suggest support for the common view that psychologists together can hardly ever agree. The 'schools' described are: functionalism, behaviourism, dynamic psychology (with a special chapter for Woodworth, Lewin and some others), hormic psychology, psychoanalysis, gestalt psychology, operationism, factor analysis and general semantics, and, cutting-up the historical sequence, and perhaps introduced principally because the author is himself interested, neoscholastic psychology.

In details the expositions could often be criticized. But this is of little account. Broadly the story is told fairly, justly and with a sense of proportion. When, as with great gusto he often does, Roback introduces his own predilection and views, it is perfectly clear that for the time being the descriptive historian is being replaced by the historical thinker or commentator. And whether the reader agrees with these interpretations and reflexions or not, they add a lot to the fun and excitement of reading.

A final chapter is called 'The Phenomenal Expansion of American Psychology'. It begins by being not very much more than a catalogue of names attached to each of the main developments of psychological teaching and research. But before long it is back to an evaluation of the present and to speculations about the future. The tone is optimistic. Roback sees American psychology now thoroughly emancipated and firmly on its feet. He envisages a future more concerned with social problems and at the same time more willing to welcome assistance from physics and chemistry, and he considers it likely that human agency itself, whether by a revival of interest in problems of volition, or by still more concern for motives and attitudes, will get enhanced recognition.

One thing is a blot on this volume. Roback seems to leave aside little opportunity of having a tilt at Boring's *History of Experimental Psychology*. This is irritating and a pity. Most of the criticisms are small ones, and in future editions it is to be hoped that they will drop out. The two books are different, in aim, in range, and in the manner of approach. They are not competitors. Probably Roback would himself be among the first to agree that Boring's great work is and will remain a classic in the field of the history of psychology.

This is a blemish. But for the rest it is a remarkably good book. No doubt there are digressions which will upset some readers; no doubt there are a number of errors in detail: these are small matters. Broadly the story that is told of the origins and development of American psychology is correct, true and stimulating. The Englishman who reads it attentively is likely to finish it with an enhanced appreciation of the great debt which we owe, and are likely in greater measure to continue to owe, to American psychology. F.C.B.

Man, Creator or Destroyer. By GEORGE MALCOLM STRATTON. London: Allen and Unwin. 1953. Pp. x+170. 16*s.*

This is a high-minded book, an expansion of two articles written in 1944 for the *Psychological Review*, but like so many high-minded books it gives but little practical guidance. It is full of 'how truths'. Man, we are told, is a creative being, and in this he differs from the animals. He creates machines, he creates works of art, he creates himself and he creates the social structures in terms of which he lives. At the same time he is capable of destruction. He can, however, forge bonds of community which take precedence over violence. He may be angry with his fellow man, but he can learn not to strike him. The violent are those whose network of attachments is limited. Most of us develop a widespread sense of respect for our neighbours, which holds us in check even when we hate them; the criminal may have a respect and affection for his nearest, but beyond that restricted circle his fellow-feeling stops. So it is with nations. Prof. Stratton contrasts the relations between the U.S.A., Great Britain, France and the Scandinavian countries with that subsisting between other nations. In the non-belligerent group there is an over-riding sense of common interest, or community, which prevents quarrels degenerating into war. The others are restrained by no such attachments.

The remedy is clear. Man's powers must be set to work to establish the bonds of common purpose which will serve to control the impulses to violence. Within the nations the bonds must unite all, and thus domestic violence will come to an end. Between the nations such attitudes of fundamental friendship and tolerance must be established that the attempt to settle disputes by war are out of the question. How true, how very true! But when we look for such practical advice as could profitably be sent to the rulers of the world, we look in vain. We can only hope that Prof. Stratton's appeal may induce his fellow countrymen to look with greater sympathy and tolerance on the aspirations of those countries which lie beyond the boundary of their present range of attachments. One gets the feeling that in Prof. Stratton's view it is always the 'others' whose sense of community is so sadly limited. W.J.H.S.

Innovation: the Basis of Cultural Change. By H. G. BARNETT. New York: McGraw-Hill. 1953. Pp. xii+462. 55*s.* 6*d.* net.

It is perhaps ungrateful to feel disappointed by this book, for Dr Barnett, an anthropologist, writes with great appreciation of the psychological approach to his topic, and the topic is one in which psychologists have shown far too little interest. Dr Barnett's main emphasis falls on the psychological process by which a new idea is generated in the individual mind, and he analyses it in much the spirit of Spearman's statement of the eduction of relations and correlates. The analysis is very likely correct, but it refers only to a small part of the total process of creative thinking. The more interesting part of the book, 'Acceptors and Rejectors', treats of the interactions that occur during social response to an innovation; and here in particular Dr Barnett brings together many apt illustrations from American life and from several primitive cultures. His points tend to be enumerated rather than organized, and are in the main rather familiar.

The psychologist's disappointment is likely to arise from the absence of fresh illumination or closer analysis of the social processes that take place when a group is faced with the invitation to change. Nevertheless, the book performs a useful service by decisively drawing attention to the topic and providing a preliminary survey for students and future inquirers in this field.

Broadmoor: A History of Criminal Lunacy and its Problems. By RALPH PARTRIDGE. London: Chatto and Windus. 1953. Pp. viii+278. 21*s.*

This is a quite unusual style of book, and one which deals with an unusual or little-known subject. The author puts into his final chapter or conclusion the kind of comments that more usually appear in a preface, for it is in this final chapter that the author states he has no claim to any special qualification in the field of criminal lunacy, and that he writes as a layman for laymen. He has, however, produced a work which can profitably be read by any lawyer, and which certainly should be read by every psychiatrist.

The thoughtful reader will find it most refreshing as a truly objective and factual presentation of a subject which more commonly is only touched upon in sensational press articles. The gruesome aspects of the subject are relieved by many touches of humour, and the whole is treated with a delicacy which does not obscure that detail which attracts the professional eye.

The author rightly pays tribute to Dr J. Stanley Hopgood, under whose superintendency Broadmoor virtually was changed from a prison into a mental hospital in very recent years. The reader who objects to the modern humanitarian and medical orientation of policy in the administration of Broadmoor will find it hard to refute its wisdom in the face of the factual account of its results with which he is presented.

OTHO FITZGERALD

Ship Without Sails. By M. Lloyd Turner. London: University of London Press. 1953. Pp. 160. 7*s.* 6*d.*

One of the greatest problems in urban life today is presented by those young people who have energy and talents for rich living but who use this energy and these talents in ways which are wasteful and harmful to the community and which do not yield for the individuals concerned any measure of real happiness. Mr Lloyd Turner describes an experiment made in East London to provide a satisfying club for boys who could find social adjustment only through participation in a small group but who, because of their own personalities and the pattern of society, had never been able to remain in ordinary clubs. The 'club rooms' were a barge moored on a Thames site, and Mr Lloyd Jones recruited members, stimulated them to make the barge a comfortable headquarters and guided the recruits through the difficult process of changing from unhelpful, uncontrolled, unsatisfied boys into a group able to make reasonably sustained efforts towards group aims and happier individual development. The experiment has had the valuable assistance of the Thames Police.

It is not easy to write up such an experiment in a way which will please both the general public, whose interest is needed, and social scientists. The defects in the book are possibly unavoidable when the double audience is considered. The author has done a good piece of work without bringing the barge club to full life. The aim has been to show the progress of the group by following developments in group relations, but this fails to a large extent because the individual personalities are not well described. The excellent wish of the author to avoid journalistic sensationalism has underplayed his very great skill and personal courage in the handling of the boys because it has served to tone down their toughness. Towards the end of the book the description of the experiment loses in firmness, and throughout one tends to feel that the photographic illustrations might have been more aptly chosen.

B.M.S.

Personality Tests and Assessments. By Philip E. Vernon. London: Methuen and Co. Ltd. 1953. Pp. xi + 220. 18*s.*

This book was badly needed from a British author. One would expect an objective and informative account of personality tests from Prof. Vernon.

Some of the author's material is based on his I.H.R.B. (1938) report. He states that the emphasis in the present work is on 'methods which are usable in educational or vocational guidance and selection, not on methods which are mainly of scientific interest'. The chapters cover interview technique, physical signs of personality, expressive movements, simple behaviour and cognitive tests, miniature and real life situation tests, ratings and judgements of personality, self-ratings and personality questionnaires, measurement of attitudes and interests, and projection techniques. An introductory chapter discusses the difficulties in the field, and here the author makes his own standpoint clear in his description of the trait-composite method, especially if it be used along with the group-difference approach, such as that of Eysenck. The final chapter briefly sums up the author's view of the present state of personality methods, and presents his conclusions.

Throughout the book much research evidence is reviewed in evaluating the tests which have been selected for discussion. In the main the author is concerned with the descriptive rather than the explanatory approach. The general impression of the book, however, is that it is another contribution to 'trait' psychology. The author emphasizes that he is concerned to show which of the tests (very few) are actually usable in applied psychology. He believes, too, that even if it is considered to be a narrow approach, the difficulties arising out of the lack of a generally accepted framework for personality description can only be surmounted in the meantime by making use of a small number of operationally defined composite variables. One might criticize the author's general approach from the point of view that he seems over-concerned with regarding personality 'tests' as potentially analogous to tests of intelligence and other abilities; one might question the existence of those many well validated tests of 'other abilities'.

The author explains why he devotes relatively little space to the tests mostly associated with clinical psychology, and Kelly and Fiske are used in the final chapter to round off the familiar criticisms of such tests.

Within its scope this book contains a great deal of information, and very good references.

FLORENCE MACNEILL

The Science of Mind and Brain. By J. S. Wilkie. London: Hutchinson's University Library: Psychology, ed. J. C. Flugel. 1953. Pp. viii + 160. 8*s.* 6*d.* net.

This is an attempt to present what is known of the relationship between psychological and neurological, and especially cortical, processes in a simple and straightforward manner. The book starts with a discussion of the nature of the problem, a review of the anatomy of the central nervous system, and a statement of the difficulties of isomorphic theories. It then discusses the problem of equivalence; the biological concept of field, the significance of brain mass; localization, and especially cortical localization of function; the frontal lobes; and ends with a discussion of the problem of integration and an appendix on Cybernetics. A glossary of terms is provided.

The author has succeeded in giving a readable and clear account of much of the more important work on the problems he discusses. Each chapter is well documented and a useful list of references is given. The balance of evidence is reflected as well as could be expected in a small book, and the author constantly encourages the reader to go to the original material.

The early part of the book is perhaps the least successful. In describing the anatomy of the central nervous system the author seems to be writing as an anatomist; that is to say, he starts from the spinal cord and works his way forward and up. In a book of this sort, where only $8\frac{1}{2}$ pages can be spared, a selection of anatomical detail on a functional basis seems desirable. The student who knows little brain anatomy is likely to find this chapter confusing. Presumably also largely from lack of space, the chapters in which philosophic problems of mind-body relationship, and of isomorphic theories especially, are discussed are also confusing. For example, the author gives the impression that a theory of isomorphism has explanatory value if and only if it postulates an exact geometric correspondence between excited cortical and, say, retinal areas. It is not at all clear, to the reviewer, at least, why this particular correspondence should have any virtue over any other fixed relationship, nor, indeed, why any statement of such a relationship should have any explanatory significance in and by itself.

This will be a very useful book for students, mainly because of the amount of material packed into its pages, but partly also because it is likely to force them into examining more closely some of the basic philosophic assumptions in this field.

G.C.D.

Bodily Physiology in Mental and Emotional Disorders. By Mark D. Altschule. New York: Grune and Stratton. 1953. Pp. 228. No price given.

The author is Assistant Professor of Medicine at Harvard Medical School, and this book has grown out of readings in psychosomatic medicine over a period of years.

It is primarily a reference book, and as such it should prove extremely useful. Various aspects of bodily physiology are considered chapter by chapter. Each chapter is planned in the same way. A brief historical sketch is followed by an examination of the effects of emotion on a particular physiological function in normal persons, then changes occurring in neuroses and psychoses are noted. Finally an attempt is made to integrate the various findings.

An idea of the comprehensiveness of the book, can be gained from the number of references quoted: there are 459 on circulatory function alone. Because of this wide coverage it is a pity that some work appears to have been overlooked. For example, the chapter on renal function omits reference to Trueta's demonstration of a possible renal shunt mechanism, and there is no mention of Selye. But it is the text itself which may arouse feelings of resentment among psychologists. The attitude throughout is one of hard-headed scepticism, which, while frequently refreshing and valuable, leads to statements such as the following: '. . . essays on the respiratory pattern as indicative of personality, mental imagery, or symbolic thinking leave the reader with a feeling of wondering skepticism', and '. . . There is no evidence that desiring motherhood causes pseudocyesis', although in this case the author thinks the development of a persistent corpus luteum might well lead to *secondary* psychological manifestations in the form of

maternal feelings. The hypothalamus may be the *via regia* between mind and body, but now apparently it is only a one-way street.

A reluctance to theorize on the psychological level is seen throughout the book and leads to preference for precise description of the physiological state and comparative disregard for the peculiarities of the psychological conditions. But perhaps this reflects the reluctance of many workers in this field to attempt such description. D.W.F.

Experimental Hypnosis. Edited by LESLIE M. LECRON. New York: Macmillan Co.; London: Macmillan and Co. Ltd. 1952. Pp. xviii+483. 45*s.* net.

This is yet another American volume of collected papers, this time in a branch of psychology in which objectivity and authority are to be welcomed.

The twenty chapters may be grouped under four heads: induction, effects, uses and explanation. Of the methods used to induce the hypnotic trance, the most searching description is made by M. H. Erickson, whose thesis is that the trance is the function of an interpersonal relation in which the willing co-operation of the subject is essential. There is no stereotyped method of induction and each subject demands individual treatment.

The papers on effects and manifestations contain all the well-known phenomena but special attention is given to age regression, time distortion and the possibilities of somatic control. Age regression is examined in detail but the checks mentioned by LeCron against simulation and role-playing are not entirely reliable. The chapter on age regression and projection tests should interest Rorschach specialists.

Of the uses of hypnosis the most promising seem to be in obstetrics, dentistry and, of course, psychotherapy. Training in general relaxation during pregnancy as a preparation for labour—what in this country we call the Read method—can be considerably helped by hypnotic suggestion. The efficacy of hypnosis in dentistry is well known. The misuses of hypnosis are mentioned and it would seem unhappily true that the hypnotic suggestion of a criminal act must remain a possibility. There are two interesting chapters on 'projective hypnoanalysis'.

On explanation the book is disappointing. Most authors are content with a rather crude adaptation of psycho-analytic concepts. Only J. B. Rhine suggests that hypnosis is a function of consciousness. W. D. Furneaux gives a dispositional explanation in terms of the correlations between the tendency to hypnosis and various forms of suggestibility. G. W. Williams defines hypnosis so widely it is surprising we are ever out of the trance state. A.G.

Psychiatry and Catholicism. By J. H. VAN DER VELDT and R. P. ODENWALD. New York and London: McGraw Hill, 1952. Pp. ix+433. 42*s.* 6*d.*

The subject-matter of this book, written primarily for the guidance of Roman Catholic clergy, psychologists and social workers, comes under three main headings. The first part of the book is devoted to philosophical considerations, and an evaluation of the 'total person' and his development considered from the scientific and religious standpoints. The second begins with a chapter on psychosomatic medicine and goes on to consider the various 'schools' of psychology and methods of psychotherapy; the existential school is described more sympathetically than others. While the great value of Freud's techniques and deductions is accepted, his philosophy is severely criticized, perhaps justly, but somewhat unscientifically. The third part discusses religion and psychiatry and contains advice to the clergy and psychiatric workers on the handling of psychiatric cases and in giving sexual instruction and dealing with marital problems; it includes a useful outline of the common psychotic, psychoneurotic and sexual disorders.

There are, inevitably, some defects. A greater clarification of the difference between man-made law and unvariable moral law, and of the meaning of mortal sin, would be helpful to non-Catholic readers. There is a tendency to deviate into religious sermons and lengthy theological quotations in the middle of a subject, instead of leaving them to a more appropriate place: a statement such as 'Psychoneurotics are mental borderline cases between the sane and insane' is liable to be unnecessarily alarming. These and other possible criticisms are, however, relatively slight when one considers the value of the work in promoting a clearer understanding of psychiatric problems. Where more than one Catholic view could be expressed the authors have chosen the most liberal; the advice given on 'counselling' and the approach to maladjusted people is full of sound common sense; and religion and psychiatry are shown not as alternatives but as mutually supplementary. W. A. H. STEVENSON

Understanding Stuttering. By A. B. Gottlober. New York: Grune and Stratton. 1953. Pp. x+274. $5·50.

After reading this book I do not yet know the cause of what we, in Britain, call stammering, and what Dr Gottlober, notwithstanding his title, calls blocking. Basically, Dr Gottlober seems to find its cause in lack of security or inadequate 'food love', but in substantiation of what must be quite the most overworked cliché in psychology today, he does not give the sort of case histories which we have the right to expect. He makes great play with the notion that children are blockers because parents, anxiously studying their children's early speech, seize on any incoordinations and nag and worry their children into blocking. All data and material are forced into an Adlerian framework: while some adult treatment of speech defects can quite reasonably proceed on Individual Psychology lines, to expect this particular School to provide the explanation for a symptom or condition which is perhaps more overdetermined than any, is to put more weight on it than it can possibly bear.

No psychologist would be advised to treat a stammerer on this book; speech therapists will get some useful wrinkles from it, and stammerers themselves some comfort, but no systematic aid. J.M.G.

Direct Analysis. Selected Papers. By John N. Rosen, M.D. New York: Grune and Stratton. 1953. Pp. 184. $3.75.

In this book, which is easy to handle and easy to read, Rosen takes us by direct route to what he calls direct analysis. A clinician who has personal experience of the extension of the principles of psychoanalysis to meet the needs of psychotic patients quickly gets into touch with Rosen's method and feels that the descriptions ring true. Some of the procedures may seem strange to those readers who are not actually engaged in this type of work; but even so there are theoretical implications in the work which cannot be ignored. Rosen is able to show that in thirty-seven treatments very great changes were observed in all, a fact which is significant even should the patients eventually relapse. It appears, however, that relapse may not occur and time may show that the results were of a permanent nature. This would accord with the reviewer's experience with children who pass through psychotic phases and with a certain type of management may come through to health without significant liability to psychotic breakdown at adolescence or in adult life.

In brief, the theory is that from the patient's point of view every psychotic had a 'bad' mother. This is only the reverse of the statement that the basis of mental health of the individual is laid down by ordinary good mothering at the start. Rosen forces his way in, displacing the 'bad' mother, and proclaiming himself as the 'good' mother; having arrived, he then deals with the highly dependent and hypersensitive psyche with great care, arranging for gradual withdrawal according to the patient's ability to make progress in emotional development from within outwards. In some cases the patient passes on to ordinary analysis, as a non-psychotic, perhaps returning for direct analysis on occasion. The result of all this is that whereas previously the patient had developed in reaction to the impingement of 'bad' mothering, there is now made possible a natural development arising out of the patient's personal spontaneity and creative capacity. It will be understood that to get at the 'badness' of the mother one has to communicate with the patient in a regressed state and the communication has to be in crude verbal or even physical terms. The actual mother may or may not be ill. When not in a regressed state the patient may be on good terms with the mother, and the Rosen technique does not in fact set people against their mothers.

This description is given here in the reviewer's own language rather than in that of Rosen, and the reader will enjoy getting at Rosen's own statement in his own words from this important and clearly written book. D. W. WINNICOTT

Encyclopedia of Aberrations. Edited by Edward Podolsky. New York: Philosophical Library. 1953. Pp. viii+550. $10.00.

This book gives a list in alphabetical order of terms and topics in psychopathology, some briefly defined, others described in fairly long articles. The longer contributions are by various authors and are as a rule reprinted from books or journals. The chief influence in the writing and editing is the psychotherapeutic outlook derived from Freudian psycho-analysis, though the contributors' opinions are far from being strictly Freudian. The book is useful as a means of quick reference to reasonably representative views of psychotherapists on the topics it lists, but it gives no systematic reference to authoritative sources and it is therefore more likely to be of use to the general reader than the student.